NOTICE

THIS REPRODUCTION WAS MADE FROM THE BEST AVAILABLE COPY OF WHICH A NUMBER OF PAGES WERE OF POOR REPRODUCTION QUALITY.

THIRTEENTH ANNUAL MANUAL JUNE 15-17, 1977 MASSACHUSETTS INSTITUTE OF TECHNOLOGY
THIRTEENTH ANNUAL MANUAL JUNE 15-17, 1977 MASSACHUSETTS INSTITUTE OF TECHNOLOGY
THIRTEENTH ANNUAL MANUAL JUNE 15-17, 1977 MASSACHUSETTS INSTITUTE OF TECHNOLOGY

THIRTEENTH ANNUAL MANUAL JUNE 15-17, 1977 MASSACHUSETTS INSTITUTE OF TECHNOLOGY
THIRTEENTH ANNUAL MANUAL JUNE 15-17, 1977 MASSACHUSETTS INSTITUTE OF TECHNOLOGY
IRTEENTH ANNUAL MANUAL JUNE 15-17, 1977 MASSACHUSETTS INSTITTHUTE OF TECHNOLOGY
ENTH ANNUAL MANUAL JUNE 15-17, 1977 MASSACHUSETTS INSTITTHUTIRTEE OF TECHNOLOGY
NNUAL MANUAL JUNE 15-17, 1977 MASSACHUSETTS INSTITTHUTIRTEE ENTH AOF TECHNOLOGY
NUAI JUNE 15-17, 1977 MASSACHUSETTS INSTITTHUTIRTEE ENTE ACFI MANNUA TECHNOLOGY
15-17, 1977 MASS L JUNE ECHNOLOGY
MASSACHUSETTS IN 17, 1977 HNOLOGY
INSTITTHUTIRTEE Proceedings USETTS MASSCIOGY
ENTH AOFI MANNUA INSTITTHUTIROGY
NUAL JUNE EC15-1 THIRTEENTH ANNUAL CONFERENCE MANNUA TENTH AY
NUAI JUNE EC15-1 ON MANNUA TENTH AY
AL JUNE EC15-17, MANUAL CONTROL MANNUA TENTH AY
UNE EC15-17, 197 JMANNUA TENTH AY
15-17, 1977 HNAC ECNNUA TENTH AY
977 ENACHUSETTS June 15-17, 1977 115-1UA TENTH AY
SETTS MASSCLTEE HNACHU TENTH AY
TEE INSTITTHUTIR MASSACHUSETTS INSTITUTE OF TECHNOLOGY TS MASSOIENTH AY
IROGOFNUL AL JMA TITTHUTTEE TH AY
UNF ECNN7, 115-1 Cambridge RCGCFNUL AL AY
977 ENACHU TSETT Massachusetts 02139 , 115-1UAUNE ECY
977 ENACHU TSETT , 115-1UAUNE ECY
7 HNACHU TSETTS 115-1UAUNE ECY
ACHU TSETTS MASSOIENINSTITTHUTTEE TH JMAIRCGCFNUI AL ANN977,7 HN 115-1UAUNE ECY
SETTS MASSCLENINSTITTHUTTEE TH JMAIRCGCFNUL AL ANN977,7 EN 1ACHU T15S MASETT-1UAUNE ECY
SSOIENINSTITTHUTTEE TH JMAIRCGCFNUI AL ANN977,7 EN 1ACHU T15S MASETT-1UAUNE ECY
ITTHUTTEE TH JMAIRCGCFNUI AL ANN977,7 HN 1ACHU T15S MASETT-1SSOIENINSTUAUNE ECY
JMAIRCGCFNUL AL ANN977,7 EN 1ACHU T15S MASETT-1SSOIENINSTUAITTHUTTEE THUNE ECY
AI ANN977,7 EN 1ACHU T15S MASETT-1SSOIENINSTUAITTHUTTEE THUNIROGOFNUL JMAE ECY
ACHU T15S MASETT-1SSOIENINSTUAITTHUTTEE THUNIROGOFNUL JMAE EN 1AI ANN977,7 ECY
SSOIENINST (NASA-CR-158107) PROCEEDINGS, 13TH ANNUAL N79-17475 :HU TY
SSOIENINST CONFERENCE ON MANUAL CONTROL (Massachusetts THRU :HU TY
OLEINSTUA Inst. of Tech.) 470 p HC A20/MF A01 N79-17520 :HU TY
INSTUAITTI CSCL 05H Unclas :HU TY
ITTHUTTEE G3/51 13767 :HU TY
E THUNIROGOFNUL JMAE EN 1AI ANN977,7 EC15SSS OLENMAINSTUASEUTTEITTHTT-1ACHU TY
CFNUL JMAE EN 1AI ANN977,7 EC15SSS CIENMAINSTUASEUTTEITTETTE THUNIROG-1ACHU TY
EN 1AI ANN977,7 EC15SSS OIENMAINSTUASEUTTEITTHTTE THUNIROG-1CF UI JMAE ACHU TY
7 EC15SSS OLENMAINSTUASEUTTEITTHTTE THUNIROG-1OFNUL JMAE ACAL ANN977,EN 1HU TY
INSTUASEUTTEITTHTTE THUNIROG-1CFNUI JMAE ACAI ANN977,EN 1HUENMA7 EC15SSS OI TY
E THUNIROG-1OFNUI JMAE ACAI ANN977,EN 1HUENMA7 EC15SSS OL TSEUTTEITTHITINSTUAY
E THUNIRCG-1OFNUI JMAE ACAL ANN977,EN 1HUENMA7 EC15SSS CI TSEUTTEITTHTTINSTUAY
THUNIROG-1CFNUL JMAE ACAI ANN977,EN 1HUENMA7 EC15SSS OI TSEE UTTEITTHTTINSTUAY
IROG-1OFNUL JMAE ACAL ANN977,EN 1HUENMA7 EC15SSS OI TSEE UTTHUNTEITTHTTINSTUAY
CFNUL JMAE ACAL ANN977,EN 1HUENMA7 EC15SSS CI TSEE UTTHUNTEIROG-1ITTHTTINSTUAY
MAE ACAL ANN977,EN 1HUENMA7 EC15SSS OI TSEE UTTHUNTEIROG-1III JCFNUTHTTINSTUAY
NN977,EN 1HUENMA7 EC15SSS OL TSEE UTTHUNTEIROG-1ITL JCFNUTHMAE ACAL ATTINSTUAY
ENMA7 EC15SSS CI TSEE UTTEUNTEIRCG-1ITL JOFNUTHMAE ACAL ATTNN977,EN 1HUINSTUAY
OI TSEE UTTHUNTEIROG-1ITI JOFNUTHMAE ACAI ATTNN977,EN 1HUIN7 EC15SSS ENMASTUAY
IROG-1ITL JOFNUTHMAE ACAI ATTNN977,EN 1HUIN7 EC15SSS ENMASTUNTECI TSEE UTTHUAY
MAE ACA OG-1Y
MAE ACA Sponsored by: National Aeronautics and Space Administration OG-1Y
E ACAI and United States Department of Transportation OC-1Y
AL ATTN CG-1Y
NN977,EN 1HUIN7 EC15SSS ENMASTUNTEOL TSEE UTTHUAITMAL E AC JAI ATTCFNUTHIROG-1Y
1HUIN7 EC15SSS ENMASTUNTECI TSEE UTTHUAITMAI E AC JAL ATTCF7,HNNNS7NUTHIRCG-1Y
15SSS ENMASTUNTECI TSEE UTTHUAITMAL E AC JAI ATTCF7,HNNNS7NU 1HUIN7 ECTHIROG-1Y

FOREWARD

This volume contains the proceedings of the Thirteenth Annual Conference on Manual Control held at the Massachusetts Institute of Technology at Cambridge from June 15 to 17, 1977. This report contains complete manuscripts of most of the papers presented at the meeting.

This was the thirteenth in a series of conferences dating back to December 1964. These earlier meetings and their proceedings are listed below:

First Annual NASA-University Conference on Manual Control, University of Michigan, December 1964. (Proceedings not printed.)

Second Annual NASA-University Conference on Manual Control, Massachusetts Institute of Technology, February 28 to March 2, 1966, NASA SP-128.

Third Annual NASA-University Conference on Manual Control, University of Southern California, March 1-3, 1967, NASA SP-144.

Fourth Annual NASA-University Conference on Manual Control, University of Michigan, March 21-23, 1968, NASA SP-192.

Fifth Annual NASA-University Conference on Manual Control, Massachusetts Institute of Technology, March 27-29, 1969, NASA SP-215.

Sixth Annual Conference on Manual Control, Wright-Patterson AFB, April 7-9, 1970.

Seventh Annual Conference on Manual Control, University of Southern California, June 2-4, 1971, NASA SP-281.

Eighth Annual Conference on Manual Control, University of Michigan, Ann Arbor, Michigan, May 17-19, 1972.

Ninth Annual Conference on Manual Control, Massachusetts Institute of Technology, May 23-25, 1973.

Tenth Annual Conference on Manual Control, Wright-Patterson AFB, April 9-11, 1974.

Eleventh Annual Conference on Manual Control, NASA-Ames Research Center, May 21-23, 1975, NASA, TM X-62,464.

Twelfth Annual Conference on Manual Control, University of Illinois and Ames Research Center, May 25-27, 1976, NASA, TM X-73,170.

CONFERENCE CHAIRMEN

Sheldon Baron
Bolt Beranek and Newman Inc.

Hendrik G. Stassen
Massachusetts Institute of Technology/
Delft University of Technology

CONTENTS

Session I: TRACKING

Chairman: D. L. Kleinman

Session II: PERFORMANCE, ATTENTION ALLOCATION AND MENTAL LOAD

Chairman: W. H. Levison

Session V: MANIPULATORS AND PROSTHETICS

Chairman: J. W. Hill

Session VI: AEROSPACE VEHICLE CONTROL

Chairman: R. E. Curry

Session VII: MOTION AND VISUAL CUES

Chairman: L. R. Young

Session VIII: DISPLAYS AND CONTROLS

Chairman: R. W. Pew

Session I

TRACKING

Chairman: D. L. Kleinman

THE OPTIMAL CONTROL FREQUENCY RESPONSE PROBLEM IN MANUAL CONTROL

By Captain Walter W. Harrington

AFFDL/FGD
Wright-Patterson AFB

SUMMARY

An optimal control frequency response problem is defined within the context of the optimal pilot model. The problem is designed to specify pilot model control frequencies reflective of important aircraft system properties, such as control feel system dynamics, airframe dynamics, and gust environment, as well as man-machine properties, such as task and attention allocation. This is accomplished by determining a bounded set of control frequencies which minimize the total control cost. The bounds are given by zero and the neuromuscular control frequency response for each control actuator. This approach is fully adaptive, i.e., does not depend upon user entered estimates. An algorithm is developed to solve this optimal control frequency response problem. The algorithm is then applied to an attitude hold task for a bare airframe fighter aircraft case with interesting dynamic properties.

INTRODUCTION

Application of the optimal pilot model to complex aircraft systems and real world tasks has identified deficiencies in the control frequency response specification [28]. Existing methods rely on user supplied estimates, either directly with respect to the control filter cutoff frequencies or time constants, or indirectly with respect to control amplitude and rate penalties [3-20,30]. Reference [28] provides some insight into the dependency of control frequency response upon aircraft system dynamics for a particular application. But, in general, there exists limited mathematical guidelines for control frequency response specification.

This paper presents a method for optimal pilot model control frequency response specification which is suitable for complex aircraft systems and tasks. The method is predictive and reflects important manned aircraft system properties, such as control feel system dynamics, airframe dynamics, gust environment, task, and attention allocation.

SYMBOLS

A_a	Augmented open loop dynamics matrix ($n_s \times n_s$)
A_p	Augmented open loop dynamics matrix containing control filter ($n_s \times n_s$)
$\bar{A}$	Augmented closed loop dynamics matrix ($n_s \times n_s$)
B_a	Augmented control distribution matrix ($n_s \times n_c$)
C_a	Augmented measurement distribution matrix ($n_m \times n_s$)
E	Expected value
E_a	Augmented disturbance distribution matrix ($n_s \times n_d$)
E_f	Filtering error matrix ($n_s \times n_s$)
E_p	Prediction error matrix ($n_s \times n_s$)
F_a	Augmented feedback matrix ($n_c \times n_s$)
J	Control cost
L	Effective feedback matrix to unaugmented state system ($n_c \times n_x$)
n_c	Number of controls
n_d	Number of disturbances
n_m	Number of measurements
n_s	Number of states in augmented system
n_x	Number of states in unaugmented system
P_a	Riccati control gain matrix for augmented system ($n_s \times n_s$)

Q	Measurement penalty matrix ($n_m \times n_m$)
R	Control rate penalty matrix ($n_c \times n_c$)
$\underline{u}$	Pilot's control input, a vector of dimension n_c
$\underline{V}_m$	Autocovariance of motor noise, a vector of dimension n_c
$\underline{V}_y$	Autocovariance of measurement noise, a vector of dimension n_m
$\underline{w}$	A disturbance vector of Gaussian white noise, a vector of dimension n_d
$\underline{x}_a$	State of the augmented system, a vector of dimension n_a
X_a	State covariance matrix of the augmented system ($n_a \times n_a$)
$\underline{y}$	A vector of measurements available to the pilot, of dimension n_m
β	Step size
Γ	Transformation matrix
Σ	Riccati filter covariance matrix
τ	Pure time delay
ω_c	Control cutoff frequency
ω_N	Neuromuscular cutoff frequency
Ω_c	Control filter matrix

Subscripts

a	Augmented
p	Perceived

Superscripts

T	Transpose
*	Optimal
^	Estimated parameter

OPTIMAL PILOT MODEL

The optimal pilot model concept, developed by Kleinman, Baron, and Levison [3-20], has demonstrated success in modeling complex, time varying control tasks. The optimal pilot model is a mathematical construct designed to synthesize pilot control performance and behavior. The model is based on the assumption that the human operator will control a dynamic, stochastic system optimally subject to his inherent limitations. These limitations are considered to be

1. A time delay, representing cognitive, visual central processing, and neuromotor delays.

2. "Remnant" signals, divided into an observation noise to represent signal degradation due to work load, scanning effects, and signal thresholds, and a motor noise to represent random errors in executing the intended control.

3. A "neuromuscular lag" to represent neuromuscular dynamics.

The control commands are synthesized by a continuous linear equalization network which contains a full state optimal filter (Kalman filter), a full state optimal predictor, and a full state optimal feedback control law. The control law is derived for an augmented state system which results from introducing the neuromuscular lag by means of a control rate penalty. The structure of the model results from a suboptimal solution to a control problem involving a time delay and observation noise. The model is shown in Figure 1.

The mathematical algorithm of the optimal pilot model is developed from the following control problem:

Given the quadratic cost functional of the form

$$J = 1/2 \int_0^\infty E\,(\underline{y}^T(t)\, Q\, \underline{y}(t) + \underline{u}_a^T(t) R \underline{u}_a(t) \qquad (1)$$

Subject to the constraints

$$\dot{\underline{x}}_a(t) = A_a \underline{x}_a(t) + B_a \underline{u}_a(t) + E_a \underline{w}_a(t) \quad (2)$$

$$\underline{y}(t) = C_a \underline{x}_a(t - \tau) + \underline{v}_y(t - \tau), \quad (3)$$

determine the non-anticipative feedback control $\underline{u}_a^*(t)$ which minimizes the cost functional.

CONTROL FREQUENCY RESPONSE IN THE OPTIMAL PILOT MODEL

The pilot control frequency response is regulated in the optimal pilot model by a first order filter matrix which processes the commanded control signals such that

$$\dot{\underline{u}} = -\Omega_c \underline{u} - \Omega_c \underline{u} \quad (4)$$

The filter matrix, Ω_c, is derived from the augmented contro' Riccati solution

$$A_a^T P_a + P_a A_a + C_a^T Q C_a - P_a B_a R^{-1} B_a^T P_a = 0 \quad (5)$$

The closed loop dynamics matrix is then given by

$$\bar{A} = A_a - B_a F_a \quad (6)$$

where the augmented feedback matrix F_a is given by

$$F_a = R^{-1} B_a^T P_a \quad (7)$$

The *augmented feedback* matrix F_a can be partitioned such that

$$F_a = [\Omega_c L | \Omega_c] \quad (8)$$

Figure 1 Structure of Optimal Pilot Model

where the filter matrix Ω_c is given by

$$\Omega_c = R^{-1} P_f \tag{9}$$

and

$$P_a = \begin{bmatrix} P & P_e^T \\ P_e & P_f \end{bmatrix} \tag{10}$$

The optimal control $\underline{u}_a^*$ is then given by

$$\underline{u}_a^* = -F_a \underline{x}_a$$

which yields equation (4)

$$\underline{u}_a^* = \underline{\dot{u}} = -\Omega_c L \underline{x} - \Omega_c \underline{u}$$

The diagonal elements ω_{c_i}, i=1, ..., n_c of Ω_c represent the first order cut-off frequencies of the control inputs u_i, i=1, ..., n_c. This can be shown by rearranging equation (4) to isolate the ith control input:

$$u_i = \frac{\omega_{c_i}}{s + \omega_{c_i}} \left[\frac{1}{\omega_{c_i}} \left(-\sum_{j=1}^{n_x} (\Omega_c L)_{ij} x_j - \sum_{\substack{j=1 \\ j \neq i}}^{n_c} \Omega_{c_{ij}} u_j \right) \right] \tag{11}$$

The cut-off frequencies are constrained such that

$$0 \leq \omega_{c_i} \leq \omega_{N_i}, \quad i=1, \ldots, n_c \tag{12}$$

where $\omega_{c_i} \geq 0$ by definition and ω_{N_i} is the pilot's neuromuscular frequency limit for the ith control input.

An iterative technique has been developed by the author [31] by which the control solution can be regulated so that the cutoff frequencies ω_{c_i}, i=1, ..., n_c, attain any desired set of values, subject to the constraints of equation (12). A technique is then required by which these cutoff frequencies be specified.

CONTROL FREQUENCY RESPONSE REQUIREMENTS

The importance of proper specification of the cutoff frequencies ω_{c_i}, i=1, ..., n_c, must be emphasized. The accuracy of the specified cutoff frequencies can greatly impact the fidelity of the optimal pilot model. Flight and simulation data [1,2] indicate real world pilot control frequency behavior not amenable to a priori estimation by existing modeling methods. User supplied estimates can thus be grossly inaccurate. In addition, these estimates impact other elements of the optimal pilot model, including the covariance propagation and optimal attention allocation algorithms. Thus the entire spectrum of performance predictions generated by the optimal pilot model depend upon the control frequency response specification.

The utility of the optimal pilot model depends, for many applications, on its predictive and adaptive capabilities. The capability to predict pilot control frequency response reflective of important manned aircraft system properties, and consistent with the other model performance predictions, would thus greatly enhance the utility of the model.

OPTIMAL CONTROL FREQUENCY RESPONSE

It is hypothesized that the human operator will adapt his control frequency response to control a dynamic, stochastic system optimally subject to his inherent limitations. These limitations include "neuromuscular lags" to represent neuromuscular dynamics. This hypothesis is a simple extension of the formulation of the optimal pilot model [3-20] as well as recent developments such as optimal attention allocation [30,32].

It is furthermore hypothesized that the control frequency response minimizes the same quadratic cost function from which the optimal pilot model is developed:

$$J = 1/2 \int_0^{\infty} E \{ \underline{y}^T(t) Q \underline{y}(t) + \underline{u}_a^T(t) R \underline{u}_a(t) \} dt \quad (1)$$

Thus it is required to determine the cutoff frequencies ω_{c_i}, i=1, ..., n_c, which minimize the cost functional. The cutoff frequencies are bounded such that

$$0 \leq \omega_{c_i} \leq \omega_{N_i}, \quad i=1, \ldots, n_c \quad (12)$$

where $\omega_{c_i} \geq 0$ by definition and ω_{N_i} is the human operator's neuromuscular frequency limit for the ith control input.

OPTIMAL CONTROL FREQUENCY RESPONSE ALGORITHM

Quadratic Cost Functional [33]

The cost functional J can be rewritten in the steady state as

$$J = E\{\underline{y}^T(t) Q \underline{y}(t) + \underline{u}_a^T(t) R \underline{u}_a(t)\} \quad (13)$$

or

$$J = E\{\underline{y}^T(t) Q \underline{y}(t) + \dot{\underline{u}}^T(t) R \dot{\underline{u}}(t)\} \quad (14)$$

This may be further rewritten as

$$J = tr[\underline{y}^T(t) Q\underline{y}(t) + R E(\dot{\underline{u}}^T(t) \dot{\underline{u}}(t))] \quad (15)$$

The control rate, $\dot{\underline{u}}(t)$, may be approximated by the pilot's own estimate of $\hat{\underline{u}}(t)$. Note that the actual $\dot{\underline{u}}(t)$ is modeled to contain white motor noises. Thus

$$\begin{aligned} \dot{\underline{u}}(t) \cong \hat{\dot{\underline{u}}}(t) &= -\Omega_c L \hat{\underline{x}}(t) - \Omega_c \underline{u}(t) \\ &= -\Omega_c L \hat{\underline{x}}(t) - \Omega_c \hat{\underline{u}}(t) - \Omega_c \underline{u}_e(t) \\ &= -F \hat{\underline{x}}_a(t) - \Omega_c \underline{u}_e(t) \end{aligned} \quad (16)$$

where $\underline{u}_e(t)$ is the error $(\underline{u}(t) - \hat{\underline{u}}(t))$. Since $\hat{\underline{x}}_a(t)$ and $\underline{u}_e(t)$ are uncorrelated for linear optimal estimation,

$$E(\hat{\underline{u}}(t)\, \hat{\underline{u}}^T(t)) = F \hat{X}_a F^T + [0 \vdots \Omega_c] (E_f + E_p) \begin{bmatrix} 0 \\ \Omega_c^T \end{bmatrix} \quad (17)$$

where E_f, E_p and $\hat{X}_a$ are the filtering error, prediction error and augmented state estimate covariance matrices, respectively. The cost function J can therefore be given by

$$J = tr[\underline{y}^T(t) Q \underline{y}(t) + R(F \hat{X}_a F^T + [0 \vdots \Omega_c](E_f + E_p) \begin{bmatrix} 0 \\ \Omega_c^T \end{bmatrix})] \quad (18)$$

Gradient Expressions

It is apparent from equation (18) and equations (4) through (10) that the cost functional J has a very complex and indirect relationship to the control cutoff frequencies ω_{c_i}, i=1, ..., n_c. Since it is unlikely that a closed-form solution for the optimal control cutoff frequencies can be found, the optimization process will be performed via a gradient algorithm similar to that developed by Kleinman [32]. Expressions will be developed for the gradient vector

$$g_{\omega_c} = \frac{\partial J}{\partial \omega_c}$$

that will be used in the subsequent optimization algorithm. Since the number of controls is usually small, this gradient vector can be evaluated numerically without incurring excessive computational cost.

This is accomplished by modifying the gradient g_{ω_c} whenever

$$g_{\omega_{c_i}} > 0 \quad \text{and} \quad \delta > \omega_{c_i} > 0 \tag{23}$$

or

$$g_{\omega_{c_i}} < 0 \quad \text{and} \quad (\omega_{N_i} - \delta) < \omega_{c_i} < \omega_{N_i} \tag{24}$$

where δ is some small frequency. Since moving in a direction opposite to the gradient would result in either a negative or a physically unattainable control cutoff frequency, the only feasible direction is given by $g_{\omega_{c_i}} = 0$. Thus ω_{c_i} remains fixed for the next iteration. The resulting gradient vector is given by $g^P_{\omega_c}$, the projected gradient vector.

Control Frequency Response Optimization

The optimization scheme to minimize the quadratic cost functional J is developed as follows. A small change in the control cutoff frequency vector along the projected gradient vector such that

$$\omega_c(k+1) = \omega_c(k) + \Delta\omega_c(k) \tag{25}$$

will still satisfy the constraints and will cause a small change in J(k),

$$J(k+1) = J(k) + (g^{P}_{\omega_c})^T \Delta\omega_c(k) \tag{26}$$

If $\Delta\omega_c(k)$ is selected as

$$\Delta\omega_c(k) = -\frac{\varepsilon\, J(k)\, g^P_{\omega_c}}{\| g^P_{\omega_c} \|^2}, \quad \varepsilon < 1 \tag{27}$$

then

$$J(k+1) = (1-\varepsilon)J(k) \tag{28}$$

Thus, each successive iteration will result in a lower cost of 100ε% if ε is sufficiently small.

A step size of the form

$$\varepsilon = \beta\, \varepsilon_{MAX}, \ \beta \le 1 \tag{29}$$

will reduce the step size when near the optimum and still satisfy the constraints if ε is sufficiently small so that

$$0 < \omega_{c_i}(k+1) < \omega_{N_i}, \ i=1, \ldots, n_c \tag{12a}$$

This can be assured by defining ε_{MAX} to be the greatest value less than unity for which

$$0 \le \left(\omega_{c_i}(k) - \frac{\varepsilon_{MAX} J(k)}{\| g_{\omega_c} \|^2}\, g_{\omega_c}\right) < \omega_{N_i} \tag{30}$$

for all i, i=1, ..., n_c. If the step size causes an increase in J, a smaller step size can be taken by reducing β. Convergence occurs when J(k+1) is sufficiently close to J(k). Note that $\omega_c(k+1)$ will continue to satisfy the constraints imposed on $\omega_c(k)$. It is only required that the user supplied initial estimate satisfy the constraints.

Computation Requirements

The optimal control frequency response gradient algorithm exercises major blocks of the basic optimal pilot model. Several considerations allow for efficient execution of this algorithm. The algorithm requires n_c (usually only one or two) attaches to a control frequency response regulation algorithm [31] to determine the transformation matrix Γ.

A frequency step size of .5 radians/second is recommended to identify appropriate control rate penalty matrix step sizes for the subsequent determination of the gradient $\frac{\partial J}{\partial r}$. This gradient requires n_c control solutions with the specified control rate penalty matrix steps, as well as n_c performance and total cost predictions. It has been found suitable for these computations to maintain constant attention allocation, thus significantly reducing the computation time which would be required by an optimal attention allocation algorithm.

The optimal control frequency response optimization algorithm forms an outer loop about the basic optimal pilot model. A complete pilot model solution, including attention allocation optimization, is required for each step. However, convergence is usually achieved in only one to five steps.

The computations required by both algorithms can be performed by a single subroutine. A flow chart of the computations required is presented in Figure 2.

APPLICATION TO AN ATTITUDE HOLD TASK

The optimal control frequency response scheme is applied to an attitude hold task for a bare airframe fighter aircraft case. An attitude hold task for a bare airframe provides a simple example which can be easily duplicated for further research in this area. The bare airframe fighter aircraft case selected has unstable longitudinal dynamics and more conventional lateral dynamics, which will be useful for the illustration of basic properties of the optimal control frequency response scheme.

Aircraft System

The fighter aircraft case involves straight and level flight at an altitude of 3,048 meters (10,000 feet) at an airspeed of 262 meters/second (862 feet/second). The airframe dynamics for this case are modeled by standard, linearized, primed, longitudinal and lateral body axis equations of motion [31]. The stability derivatives and other parameters pertinent to this case are presented in Table 1. For simplicity, as well as to accentuate the aircraft dynamics, no models are included for the stability augmentation system or control feel system. The airframe is disturbed by turbulence with gust intensities of 5 feet/second. MIL SPEC 8785B turbulence is provided, as modeled by Heath [24].

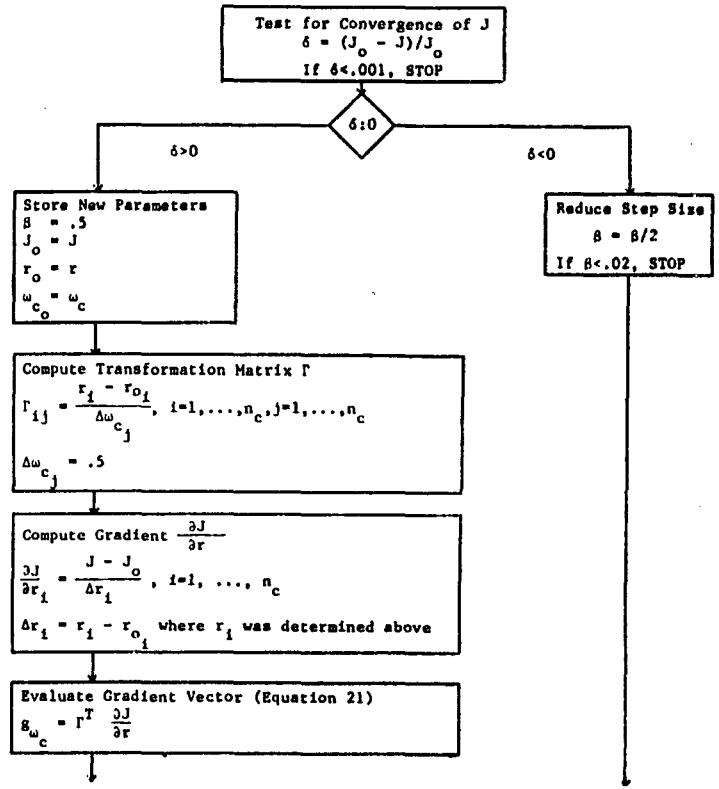


FIGURE 2. OPTIMAL CONTROL FREQUENCY RESPONSE OPTIMIZATION ALGORITHM FLOW CHART (1 of 2)

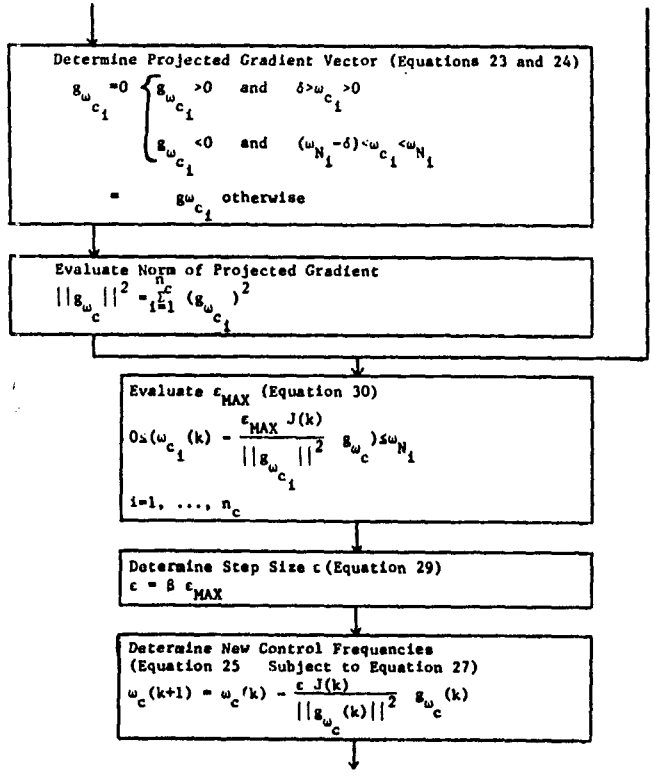


FIGURE 2. OPTIMAL CONTROL FREQUENCY RESPONSE OPTIMIZATION ALGORITHM FLOW CHART (2 of 2)

PARAMETER	LONGITUDINAL AXIS SYSTEM	LATERAL AXIS SYSTEM
Unprimed Stability Derivatives	X_u = -.01578 X_w = .02144 $X_{\delta e}$ = 20.6 M_u = -.001169 M_w = .005766 M_q = -.6528 $M_{\delta e}$ = -28.2 $M_{\dot{w}}$ = -.0003094 Z_u = -.06086 Z_w = -1.692 $Z_{\delta e}$ = -157.3	Y_v = -.2932 $Y_{\delta a}$ = 14.15 $Y_{\delta r}$ = 50.72 L_β = -70.95 L_p = -2.551 L_r = -.06964 $L_{\delta a}$ = -62.92 $L_{\delta r}$ = 17.01 N_β = 11.51 N_p = -.01704 N_r = -.2266 $N_{\delta a}$ = -2.261 $N_{\delta r}$ = -9.057
Moments		I_{xx} = 9017.5 I_{zz} = 58104. I_{xz} = 144.93
Wing Span	30.	30.
Total Velocity	862.0	862.0
Flight Path Angle	0.	0.
Angle of Attack	1.5285	1.5285
Altitude	10000.	10000.
Units	Feet, Seconds, Radians	Feet, Seconds, Radians

TABLE 1. FIGHTER AIRCRAFT STABILITY DERIVATIVES AND OTHER PARAMETERS

Controls include a side stick for pitch and roll and pedals for yaw. The neuromuscular frequency response limits are estimated to be 10, 10, and 5 radians/second for the pitch stick, roll stick, and pedals, respectively [25,26]. The multi-dimensional control for the lateral axis system will be useful to demonstrate the utility of the optimal control frequency response algorithms.

Measured quantities include pitch angle, vertical velocity, roll angle, and yaw angle. It is assumed that the rates of these quantities are simultaneously available. The perceptual thresholds for the measured quantities are presented in Table 2.

Task

The task is defined to be precision attitude hold in turbulence. Control action is required as soon as a deviation in measured quantities is noticed, and deviations must be maintained within specified limits. A total of 80% attention is allocated to this task.

Immediate control action implies that the indifference thresholds are the same as the perceptual thresholds, as presented in Table 2. The maximum allowable deviations are also presented in this table. The optimal pilot model state penalties are derived by the conventional rule

$$Q_{ii} = \left(\frac{1}{\text{MAX}|y_i(t)|}\right)^2 \qquad (31)$$

and are presented in Table 2.

Model Predictions

The optimal control frequency response scheme is implemented on the AFFDL optimal pilot model computer program FGDPILOT [31], which is designed for complex aircraft system and simulation analysis. The program solves the longitudinal and lateral axis systems separately, which results in a significant cost reduction. However, this presents a problem for specification of attention allocation between the two axis systems. For the purpose of this paper, the naive assumption will be made of equal attention allocation, i.e., 40% total attention per axis system. This problem is addressed by the related conference paper "An Approach to the Multi-Axis Problem in Manual Control" by this author [34].

MEASURED QUANTITY	UNITS	PERCEPTUAL THRESHOLD	INDIFFERENCE THRESHOLD	MAXIMUM ALLOWABLE DEVIATION	STATE PENALTY
Pitch Angle	deg.	.8	.8	10.	.01
Pitch Rate	deg./sec.	1.6	1.6	32.	.001
Vertical Velocity	ft./sec.	.27	.27	10.	.01
Vertical Acceleration	ft./sec.^{2}	.54	.54	32.	.001
Roll Angle	deg.	.9	.9	10.	.01
Roll Rate	deg./sec.	1.8	1.8	32.	.001
Yaw Angle	deg.	.9	.9	10.	.01
Yaw Rate	deg./sec.	1.8	1.8	32.	.001

TABLE 2. MEASURED QUANTITIES AND RELATED PARAMETERS

The computer program is completely automated, and requires essentially only those parameters presented in this section. The program contains nominal values for all the basic optimal pilot model parameters. Program predictions include optimal control frequency response, optimal attention allocation [32], and control and aircraft system performance. Predictions for this case are presented in Figures 3 and 4 for the longitudinal and lateral axis systems, respectively.

The predicted control frequency responses for this case are presented in Table 3. The maximum physically attainable frequency response is predicted for the longitudinal axis system, as would be expected for an unstable system. Intermediate values are predicted for the roll stick and pedal frequency response. These values would have been difficult to predict a priori. However, these predictions are based on equal attention allocation to the longitudinal and lateral axis systems. It is unlikely that equal attention allocation is appropriate for this case. Since the control frequency response predictions are dependent upon attention allocation, the predictions are in error. Thus the multi-axis attention allocation problem must be addressed.

CONTROLLER	NEUROMUSCULAR CUTOFF FREQUENCY	OPTIMAL CONTROL CUTOFF FREQUENCY
Pitch Stick	10 radians/second	10 radians/second
Roll Stick	10 radians/second	7.6 radians/second
Pedals	5 radians/second	4.4 radians/second

TABLE 3. PREDICTED CONTROL FREQUENCY RESPONSE

The multi-axis problem is addressed by reference [34], in which corrected predictions are made for this case. The multi-axis solution indicates a large shift in attention allocation to the unstable longitudinal axis system. The predicted attention allocation is .697 and .103 for the longitudinal and lateral axis systems, respectively. The corrected control frequency responses are presented in Table 4. The maximum physically attainable frequency response is again predicted for the unstable longitudinal axis system. But, a substantial reductior in control frequency response is predicted for the lateral axis system. This reduction is most consistent with the reduction in attention allocation to this axis system. Note that the pedal cutoff frequency of .2 radians/second corresponds to a control cycle time of about 30 seconds. These values could not be predicted by previous methods.

FIGURE 3. OPTIMAL PILOT MODEL PERFORMANCE PREDICTIONS FOR LONGITUDINAL F-16 BARE AIRFRAME ATTITUDE HOLD TASK (1 of 2)

CASE DESCRIPTION

LONGITUDINAL DYNAMICS, H=.8
PRECISION ATTITUDE HOLD TASK
LONGITUDINAL DYNAMICS

ALTITUDE = 1.000000E+04 FEET
NOMINAL AIRSPEED = 8.616933E+02 FEET/SECOND
NOMINAL ANGLE OF ATTACK = 1.528500E+00 DEGREES
TURBULENCE CONDITIONS
SIG U GUST = 5.000000E+00 FEET/SECOND
SIG V GUST = 5.000000E+00 FEET/SECOND
SIG W GUST = 5.000000E+00 FEET/SECOND

AIRFRAME EIGENVALUES

(-3.668128З)+J(0.)
(1.0855377)+J(0.)
(-.2229867ВЕ-01)+J(.13527864)
(-.2229867ВЕ-01)+J(-.13527864)

PILOT MODEL PARAMETERS

TOTAL ATTENTION ALLOCATION TO PRIMARY FLIGHT TASK (ATTN) = .400000E+00
PURE TIME DELAY (TAU) = .200000E+00 SECONDS
MOTOR NOISE TO SIGNAL RATIO (PM) = 3.000000E-03
MEASUREMENT NOISE TO SIGNAL RATIO (PY) = 1.000000E-02
TOTAL COST (COST) = 2.505697E+00
SCANNING COST (SCOST) = 1.790228E+00

CONTROLS

CONTROLLER	NEUROMUSCULAR FREQUENCY	OPTIMAL CONTROL FREQUENCY	CONTROL RATE PENALTY (RINV)	CONTROL FORCE STD DEVIATION	CONTROL SURFACE STD DEVIATION
PITCH STICK	10.000000	10.314665	9.8533666	.31868657	.31868657

OBSERVED QUANTITIES

OBSERVED QUANTITY	UNITS	TASK VECTOR	PERCEPTUAL THRESHOLD	INDIFFERENCE THRESHOLD	OPTIMAL ATTENTION ALLOCATION	STANDARD DEVIATION
PITCH ANGLE	DEG	.10000000E-01	.80000000	.80000000	.30604225E-01	4.7593264
PITCH RATE	DEG/SEC	.10000000E-02	1.6000000	1.6000000	.30604225E-01	1.9888165
VERTICAL VELOCITY	FT/SEC	.10000000E-01	.27000000	.27000000	.36939577	10.666126
VERTICAL ACCEL	FT/SEC**2	.10000000E-02	.54000000	.54000000	.36939577	26.346301

FIGURE 3. OPTIMAL PILOT MODEL PERFORMANCE PREDICTIONS FOR LONGITUDINAL F-16 BARE AIRFRAME ATTITUDE HOLD TASK (2 of 2)

CASE DESCRIPTION

LATERAL DYNAMICS, BARE AIRFRAME, M=.8
PRECISION CONTROL WITH ROLL STICK AND PEDALS
LATERAL DYNAMICS

ALTITUDE = 1.000000E+04 FEET
NOMINAL AIRSPEED = 8.616933E+02 FEET/SECOND
NOMINAL ANGLE OF ATTACK = 1.528500E+00 DEGREES
TURBULENCE CONDITIONS
SIG U GUST = 5.000000E+00 FEET/SECOND
SIG V GUST = 5.000000E+00 FEET/SECOND
SIG W GUST = 5.000000E+00 FEET/SECOND

AIRFRAME EIGENVALUES
(0.)+J(0.)
(-2.5341997)+J(0)
(-.18760888E-01)+J(0)
(-.25919919)+J(3.6290849)
(-.25919919)+J(-3.6290849)

PILOT MODEL PARAMETERS

TOTAL ATTENTION ALLOCATION TO PRIMARY FLIGHT TASK (ATTN) = .400000E+00
PURE TIME DELAY (TAU) = .200000E+00 SECONDS
MOTOR NOISE TO SIGNAL RATIO (FM) = 3.000000E-03
MEASUREMENT NOISE TO SIGNAL RATIO (PY) = 1.000000E-02
TOTAL COST (COST) = 1.802367E-01

FIGURE 4. OPTIMAL PILOT MODEL PERFORMANCE PREDICTIONS FOR LATERAL F-16 BARE AIRFRAME ATTITUDE HOLD TASK (1 of 2)

CONTROLS

CONTROLLER	NEUROMUSCULAR FREQUENCY	OPTIMAL CONTROL FREQUENCY	CONTROL RATE PENALTY (RINV)	CONTROL FORCE STD DEVIATION	CONTROL SURFACE STD DEVIATION
ROLL STICK	10.000000	7.5831960	168.07812	.18555362	.18555362
PEDALS	5.0000000	4.3811237	390.77780	.37499172	.37499172

OBSERVED QUANTITIES

OBSERVED QUANTITY	UNITS	TASK VECTOR	PERCEPTUAL THRESHOLD	INDIFFERENCE THRESHOLD	OPTIMAL ATTENTION ALLOCATION	STANDARD DEVIATION
ROLL ANGLE	DEG	.10000000E-01	.90000000	.90000000	.30351583	3.1144329
ROLL RATE	DEG/SEC	.10000000E-02	1.8000000	1.8000000	.30351583	7.8006640
YAW ANGLE	DEG	.10000000E-01	.90000000	.90000000	.96484172E-01	.82409736
YAW RATE	DEG/SEC	.10000000E-02	1.8000000	1.8000000	.96484172E-01	1.6599728

FIGURE 4. OPTIMAL PILOT MODEL PERFORMANCE PREDICTIONS FOR LATERAL F-16 BARE AIRFRAME ATTITUDE HOLD TASK (2 of 2)

TABLE 4. PREDICTED MULTI-AXIS CONTROL FREQUENCY RESPONSE

CONTROLLER	NEUROMUSCULAR CUTOFF FREQUENCY	MULTI-AXIS OPTIMAL CONTROL CUTOFF FREQUENCY
Pitch Stick	10.rad/sec	10.rad/sec
Roll Stick	10.rad/sec	4.2 rad/sec
Pedals	5.rad/sec	.20rad/sec

The multi-axis solution also produced a substantial reduction in total quadratic cost, as shown in Table 5. This is further evidence of the adaptive capability of the methods developed herein and in reference [34].

TABLE 5. PREDICTED QUADRATIC COST

AXIS SYSTEM	QUADRATIC COST	MULTI-AXIS QUADRATIC COST
Longitudinal	2.51	1.29
Lateral	.18	.35
TOTAL SYSTEM	2.69	1.64

SUMMARY OF THE OPTIMAL CONTROL FREQUENCY RESPONSE PROBLEM

A method has been presented for optimal pilot model control frequency response specification which is suitable for complex aircraft systems and tasks. The method is predictive and reflects important manned aircraft system properties, such as control feel system dynamics, airframe dynamics, gust environment, task, and attention allocation. Dependency upon other optimal pilot model predictions results from the definition and iterative nature of the method.

The method has been applied to an attitude hold task for a bare airframe fighter aircraft case. This application demonstrated the method's capability to make realistic predictions for stable as well as unstable aircraft system dynamics and for scalar as well as multi-dimensional controls. The application identified a related deficiency in attention allocation specification for the multi-axis control task. This problem is addressed in the conference paper "An Approach to the Multi-Axis Problem in Manual Control" by this author.

RECOMMENDED AREAS OF INVESTIGATION

Control frequency response is considered to be bounded by neuromuscular control frequency response limits. First order approximations are used in the optimal pilot model. It would be beneficial to this area of research to experimentally identify these limits for conventional controllers. Suggested controllers are force and dynamic side sticks, center sticks with high and low roll pivot, a wheel and column combination, and pedals.

Further investigations are required to fully develop the optimal control frequency response concept. Modifications to the motor noise [35,36,37] from the existing implementation may have significant effect upon the total system cost and thus the optimal control frequency response. Control cost may require additional terms other than rms control activity [36], both for control frequency response optimization and for total attention allocation specification. And, of course verification against experimental data is required.

REFERENCES

1. Gressang, Stone, Pollard, and Kugel: Low Visibility Landing Pilot Modeling Experiment and Data, Phase I. AFFDL TR 75-41, April 1976.

2. Gressang: Low Visibility Landing Pilot Modeling Experiments and Data, Phase II. AFFDL TR 75-57, August 1975.

3. McRuer and Krendel: Mathematical Models of Human Pilot Behavior. AGARDograph No. 188, January 1974.

4. Baron and Kleinman: The Human As An Optimal Controller and Information Processor. IEEE Trans. Man-Machine Systems, Vol. MMS-10, March 1969, pp 9-17.

5. Baron et al: Application of Optimal Control Theory to the Prediction of Human Performance In A Complex Task. AFFDL TR 69-81, March 1970.

6. Kleinman, Baron, and Levison: A Control Theoretic Approach to Manned-Vehicle Systems Analysis. IEEE Trans. Auto. Control, Vol. AC-16, December 1971, pp 824-832.

7. Kleinman, Baron, and Levison: An Optimal-Control Model of Human Response, Part I: Theory and Validation. Automatica, Vol. 6, 1970, pp 357-369.

8. Kleinman: Optimal Control of Linear Systems with Time-Delay and Observation Noise. IEEE Trans. Auto. Control, Vol. AC-14, October 1969, pp 524-527.

9. Graham and McRuer: Analysis of Nonlinear Control Systems. Dover Publications, New York, 1971, Chapter 6.

10. Baron and Levison: An Optimal Control Methodology for Analyzing the Effects of Display Parameters On Performance and Work Load in Manual Flight Control. IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-5, July 1975, pp 423-430.

11. Bryson and Ho: Applied Optimal Control. Ginn and Co., Waltham, Mass., 1969, Chapter 5.

12. Kleinman and Baron: Analytic Evaluation of Display Requirements for Approach to Landing. NASA CR-1952, October 1971.

13. Kleinman and Perkins: Modeling Human Performance in a Time-Varying Anti-Aircraft Tracking Loop. IEEE Trans. Auto. Control, Vol. AC-19, August 1974, pp 297-306.

14. Harvey: Application of An Optimal Control Pilot Model to Air-to-Air Combat. AFIT Thesis GA/MA/74M-1, March 1974.

15. Dillow, Picha and Anderson: Slushy Weightings for the Optimal Pilot Model. 11th Annual Conference on Manual Control, May 1975.

16. Kleinman: Computer Programs Useful in Linear Systems Studies. Systems Control, Inc., Technical Memorandum, December 1971.

17. Kleinman: On An Iterative Technique for Riccati Equation Computation. IEEE Trans. Auto. Control, Vol. AC-13, February 1968, pp 114-115.

18. Kwakernaak and Sivan: Linear Optimal Control Systems. John Wiley, New York, 1972, Chapter 3.

19. McRuer, Ashkenas, and Graham: Aircraft Dynamics and Automatic Control. Princeton University Press, 1973, Chapters 5 and 6.

20. Kleinman and Killingsworth: A Predictive Pilot Model for STOL Aircraft Landing. NASA CR-2374, March 1974.

21. Gressang: A Note on Solving Riccati Equations Associated with Optimal Pilot Models. AFIT/AIAA Mini Symposium on Recent Advances in Aeronautical Research Development, and System WPAFB, 26 March 1975.

22. Bryson and Hall: Optimal Control and Filter Synthesis by Eigenvector Decomposition. Report, Dept. of Aero. and Astro., Stanford University, December 1971.

23. McRuer and Graham: Human Pilot Dynamics in Compensatory Systems. AFFDL TR 69-72, July 1965.

24. Heath: State Variable Model of Wind Gusts. AFFDL/FGC TM-72-12, July 1972.

25. Stark: Neurological Control Systems, Studies in Bioengineering. Plenum Press, New York, 1968.

26. Milsum: Biological Control Systems Analysis. McGraw-Hill Book Company, New York, 1966.

27. Curry, Young, Hoffman, and Kugel: A Pilot Model with Visual and Motion Cues. AIAA Visual and Motion Simulation Conference, Dayton, Ohio, April 1976.

28. Harrington: The Application of Pilot Modeling to the Study of Low Visibility Landing. Twelfth Conference on Manual Control, May 1976.

29. Hess: Prediction of Pilot Opinion Ratings Using An Optimal Pilot Model. To appear in Human Factors.

30. Curry, Hoffman, Young: Pilot Modeling for Manned Simulation. AFFDL-TR-76-124, Volume I, December 1976.

31. Harrington: A Computer Program for the Analysis of Manned Aircraft and Simulation Systems. AFFDL TR to be published.

32. Kleinman: Solving the Optimal Attention Allocation Problem in Manual Control. IEEE Trans. Auto. Control, Vol. AC-21, No. 6, December 1976, pp 813-821.

33. Hoffman, Curry, Kleinman, Hollister, Young: Display/Control Requirements for VTOL Aircraft. ASI-TR-75-26 (NASA CR 145026), August 1975.

34. Harrington: An Approach to the Multi-Axis Problem in Manual Control. Thirteenth Conference on Manual Control, June 1977.

35. Kleinman and Baron: Manned Vehicle Systems Analysis by Means of Modern Control Theory. NASA CR-1753, June 1971.

36. Levison, Baron, and Junker: Modeling the Effects of Environmental Factors on Human Control and Information Processing. AMRL-TR-76-74, August 1976.

37. Levison and Junker: A Model for the Pilot's Use of Motion Cues in Roll-Axis Tracking Tasks. Thirteenth Conference on Manual Control, June 1977.

THE EFFECTS OF DEVIATE INTERNAL REPRESENTATIONS IN THE OPTIMAL MODEL OF THE HUMAN OPERATOR

by

Sheldon Baron and Jeffrey E. Berliner
Bolt Beranek and Newman Inc.
Cambridge, Massachusetts

May 1977

1. Introduction

Modern control and estimation theory have been used successfully to develop a model for human performance in continuous control tasks [1]. This model, frequently referred to as the optimal control model of the human operator, has been validated extensively by experimental data and has been applied to a variety of problems. The model incorporates an "internal model" that is an exact replica of the system model as part of a Kalman filter sub-model that represents human information processing. The concept that the human operator builds an internal model of his "universe" (e.g., through training) is not uncommon in psychology. Moreover, the assumption of a perfect internal model appears to be a satisfactory one in many instances, as has been demonstrated by the agreement between model predictions and experimental data.

There are situations, however, in which the assumption of a perfect model does not appear suitable and important applications which would benefit from allowing for an internal model that is different from the system model. For example, naive or untrained trackers may not have "perfect" models even for simpler systems. Tracking of targets executing deterministic but unknown motions requires admitting imperfect internal models (for the input) for complete generality. When a system failure occurs there is a change in the system; until this change is detected and the failed system identified the operator's model is different than the system model.

In this note, some of the issues and equations involved in predicting closed-loop man-machine performance for situations in which the human operators' knowledge of the system and/or environment are imperfect are presented and discussed. Several examples to demonstrate some of the effects to be expected when such is the case are then given. Details concerning equation development may be found in [2].

2. Equations for Deviate Internal Model

Let the system to be controlled by the human operator be described by the linear equations

$$\dot{\underline{x}}(t) = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{u}(t) + \underline{F}\,\underline{w}(t) \qquad (1)$$

$$\underline{y}(t) = \underline{C}\,\underline{x}(t) + \underline{D}\,\underline{u}(t) \qquad (2)$$

where $\underline{x}$ is an n_x dimensional vector of system state variables, $\underline{u}$ is an n_u-dimensional vector of control inputs, $\underline{y}$ is an n_y-dimensional vector of displayed outputs and $\underline{w}$ is an n_w-dimensional vector of a zero-mean, gaussian, white noise process with auto-covariance $E(\underline{w}(t_1)\underline{w}'(t_2)) = \underline{W}\,\delta(t_1-t_2)$. We assume $\underline{w}(t)$ is stationary so that $\underline{W}$ is constant for all t. We will also assume that the matrices in (1) and (2) are constant Thus, we treat a time-invariant system. Moreover, we will be concerned here only with the steady-state solution.

The optimal control model for the human operator has the structure illustrated in Figure 1. The structure and equations of Figure 1 have been documented in [1]. The blocks in Figure 1 labeled estimator and predictor model human information processing. For these processes to be performed "optimally" it is necessary to have perfect knowledge of the system ($\underline{A}$, $\underline{B}$, $\underline{C}$, $\underline{D}$, $\underline{F}$), the driving noise-statistics ($\underline{W}$), and the parameters describing human limitations (τ, $\underline{T}_n$, $\underline{V}_y$, $\underline{V}_m$). The control gains, $\underline{L}^*$, model human control-command generation or compensation and are selected so as to minimize a quadratic cost functional. To achieve a minimum, i.e., to compute $\underline{L}^*$, it is necessary to know $\underline{A}$,$\underline{B}$ and the weighting coefficients ($q(\cdot)_i$). Thus, there are three classes of quantities or parameters (system/environment, own limitations, and cost weightings) that are required to be known by the human operator if he is to perform optimally.

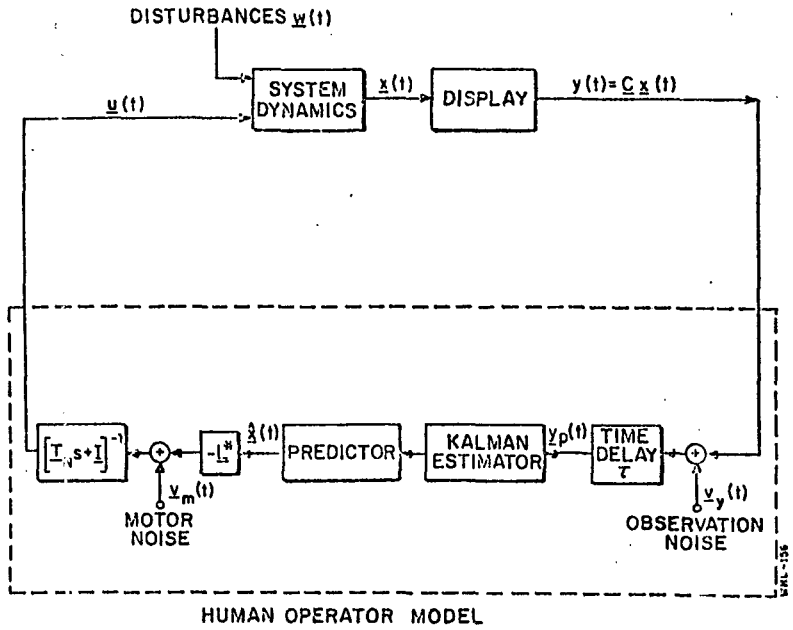


Figure 1. Structure of the Optimal Control Model

There are many assumptions that can be made concerning the human operator's knowledge of the requisite information. At one extreme, one can assume that all quantities are unknown (including the dimensions of the various matrices). At the other end of the spectrum, one can assume that all quantities are known and the human performs optimally. This latter assumption is, of course, the one used in formulating the optimal control model; for trained operators, it seems closer to the truth (or, at least it explains the data better) that the assumption of complete ignorance. Here, for reasons discussed in [2], we assume the human operator knows the cost functional weightings and his own limitations of delay, neuromotor-lag and observation noise. On the other hand, we allow the system matrices to be unknown (even as to dimension) and the motor-noise also to be unknown.

To implement the above assumptions, we assume the human operator's internal model to be

$$\dot{\underline{z}}(t) = \tilde{\underline{A}}_1 \underline{z}(t) + \tilde{\underline{B}}_1 \underline{u}_c(t) + \tilde{\underline{E}}_1 \tilde{\underline{v}}_1(t) \qquad (3)$$

$$\underline{y}(t) = \tilde{\underline{C}}_1 \underline{z}(t) \qquad (4)$$

$$E(\tilde{\underline{v}}_1(t_1)\, \tilde{\underline{v}}_1'(t_2)) = \tilde{\underline{V}}_1\, \delta(t_1 - t_2) \qquad (5)$$

where

$$\tilde{\underline{A}}_1 = \begin{bmatrix} \underline{A} & \underline{B} \\ 0 & -\underline{T}_N^{-1} \end{bmatrix}, \quad \tilde{\underline{B}}_1 = \begin{bmatrix} 0 \\ \underline{T}_N^{-1} \end{bmatrix}, \quad \tilde{\underline{C}}_1' = \begin{bmatrix} \underline{C}' \\ \underline{D}' \end{bmatrix}$$

$$\tilde{\underline{E}}_1 = \begin{bmatrix} \tilde{\underline{E}} & 0 \\ 0 & \underline{T}_N^{-1} \end{bmatrix}, \quad \tilde{\underline{V}}_1 = \begin{bmatrix} \tilde{\underline{W}} & 0 \\ 0 & \tilde{\underline{V}}_n \end{bmatrix} \qquad (6)$$

where the matrices with "tildes" indicate internal matrices and Equations (1) and (2) have been "augmented" to incorporate the "neuromotor" dynamics (see Fig. 1 and [1]). The perceived variables remain unchanged inasmuch as the "true" $\underline{y}$ is displayed to the operator. The "internal state" $\underline{z}$ does not have to be of the same dimension as $\underline{x}$. However, we assume that $\underline{y}$ and $\underline{u}$ in the

internal model have the same dimensions as the corresponding vectors of the system.

It is now assumed that the human will perform "optimally" for his internal system. These assumptions lead to a set of coupled delay-differential equations. In the special case, where $\underline{x}_j = \underline{Z}$ and $\tilde{\underline{C}}_1 = \underline{C}_1$, the following equations describing closed-loop performance are obtained [2]

$$
\begin{aligned}
\dot{\underline{e}}(t) &= (\tilde{\underline{A}}_1 - \tilde{\underline{K}}\,\underline{C}_1)\,\underline{e}(t) + (\underline{A}_1 - \tilde{\underline{A}}_1)\,\underline{x}(t-\tau) \\
&\quad + \underline{F}_1\,\underline{w}(t-\tau) - \tilde{\underline{K}}\,\underline{v}_y(t-\tau) \\
\dot{\hat{\underline{x}}}(t) &= (\tilde{\underline{A}}_1 - \underline{B}_1\,\tilde{\underline{L}})\,\hat{\underline{x}}(t) + e^{\tilde{\underline{A}}_1^T \tau}\,\tilde{\underline{K}} \\
&\quad [\underline{C}_1\,\underline{e}(t) + \underline{v}_y(t-\tau)] \\
\dot{\underline{x}}_1(t) &= \underline{A}_1\,\underline{x}_1(t) - \underline{B}_1\,\tilde{\underline{L}}\,\hat{\underline{x}}(t) + \underline{F}_1\,\underline{w}(t) \quad .
\end{aligned}
\tag{7}
$$

where e(t) is the state estimation error and $\tilde{K}$ is the Kalman gain for the system described by Equations (3)-(6). Equation (7) is also a "coupled" set of delay-differential equations. Note, however, that if $\tilde{A}_1 = A_1$ the equation for the error "decouples" from the state equation and the estimation equation. Moreover, the system reduces to a set of ordinary differential equations. Performance computations are thereby simplified enormously requiring evaluation of $n_x \times n_x$ matrices only. This is the case even if $\tilde{W}_1 \neq W_1$. Unfortunately, the assumptions required to achieve this simplification are too stringent for most considerations.

The delay-differential character of the above equations can be circumvented by approximating the human's delay via a Padé approximation. The delay is then considered part of the system dynamics (except for computation of human describing functions); it is a part that is assumed known to the human operator so there will be some compensation for the delay. The resulting closed-loop equations are linear and time invariant. However, their stability is not automatically guaranteed as in the case when all matrices are known to the operator; instead, stability depends on the particular internal model selected. The necessary modifications are given in [2].

3. Examples

Incorrect Knowledge of Vehicle Dynamics

In order to control a vehicle, the pilot must learn its basic response characteristics. One can readily envision this as a two-stage process: 1) the development of an appropriate structural model for the plant; and 2) the adjustment (or fine tuning) of the parameters in that structure. Such a process is consistent with the notions of system identification theory. With regard to structure, the problem in a multi-input, multi-output plant involves learning the couplings as well as the basic modes of response. For single-input, single-output situations a fundamental issue is the order of the plant dynamics, i.e., how many integrations are there between control input and plant output.

Figures 2 and 3 show the predicted describing function and remnant for an operator optimizing his performance based on different internal models of the vehicle dynamics. In each case, the input disturbance was filtered white noise with a 2 rad/sec bandwidth. In Figure 2, the true plant dynamics are K/s, i.e., the rate-of-change of plant output is proportional to the control input. Three curves are shown: one in which the operator has the correct model, one in which the internal model is incorrect (the output is proportional to the input), and one in which the operator has a large pseudomotor noise [2]. The curve corresponding to having the correct model agrees quite well with the measured describing functions for this case [1]. Note that the effect of having the wrong internal model is substantial whereas the effect of high pseudo motor-noise is slight (a reduction in gain at low frequencies.

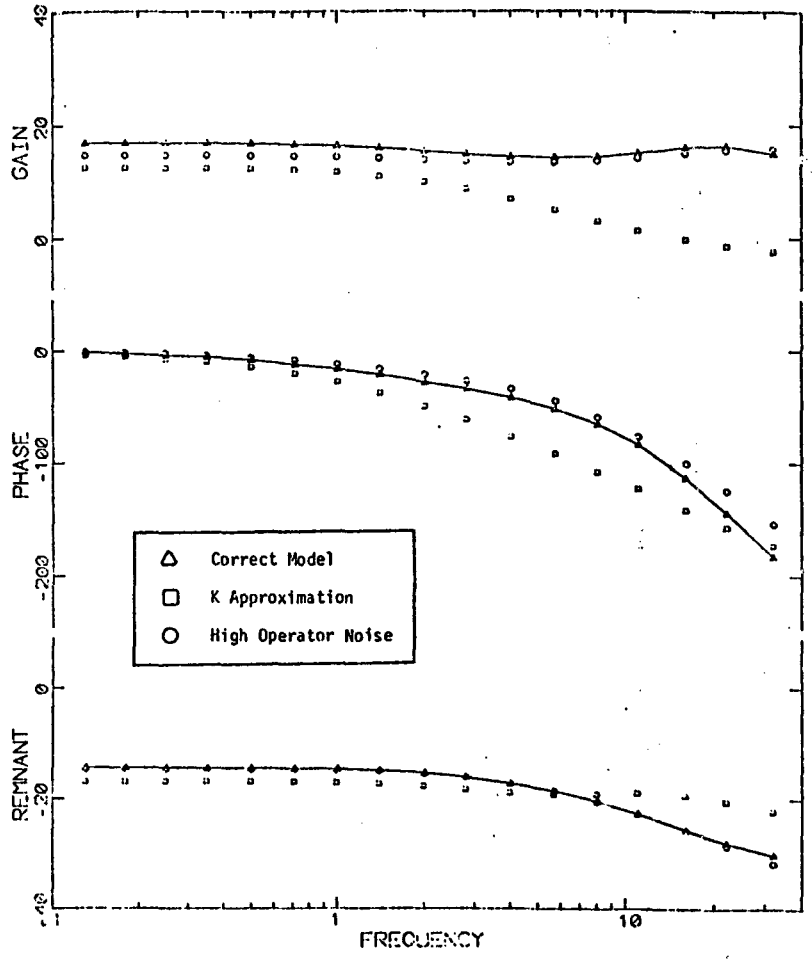


Figure 2. Effects of Deviate Internal Model and High Operator Noise on K/s Regulation

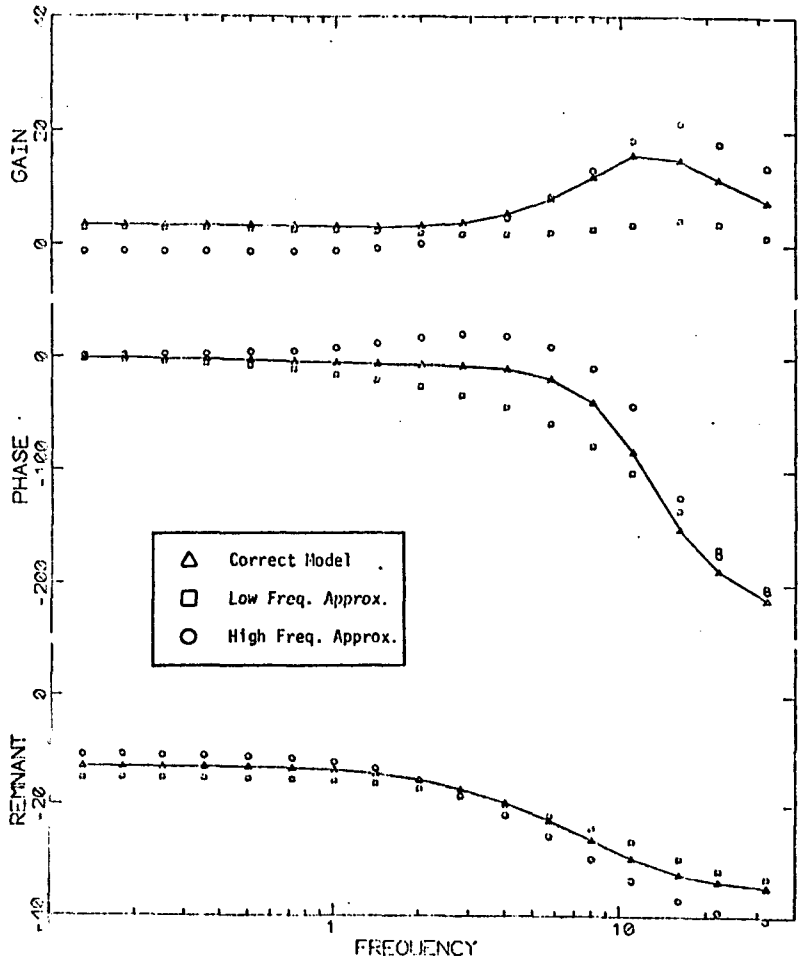


Figure 3. Effect of Deviate Internal Model on Regulation of Roll Disturbance

Figure 3 presents similar results for a more complex plant which represents the roll-synamics of an aircraft. Results are shown for the case in which the operator has the correct internal model and for the cases where the model is a good approximation to either the low frequency or high frequency plant response. Frequency characteristics of the three vehicle models are plotted in Figure 4. Again the results show that we can expect measurements of the pilot's describing functions to be significantly different if operating with different internal models. In this case, the remnant is somewhat less revealing.

Model results were also obtained for the case where the pilot's internal model of the aircraft roll dynamics differed only slightly from the actual dynamics over the entire frequency range of interest. In this case (not shown), differences in the above measurements were not distinguishable from those that might be due to othe parameter changes.

On the basis of these preliminary results, we believe that major structural errors in the operator's internal model of the plant dynamics can be inferred by comparing the measured describing function and remnant with that predicted, by the OCM, for a trained operator. Moreover, the form of the operator's internal model may also be deduced using the OCM. Major parameter errors are also probably discernible. However, the fine-tuning process of model identification may not be distinguished readily from other factors such as general noise-reduction.

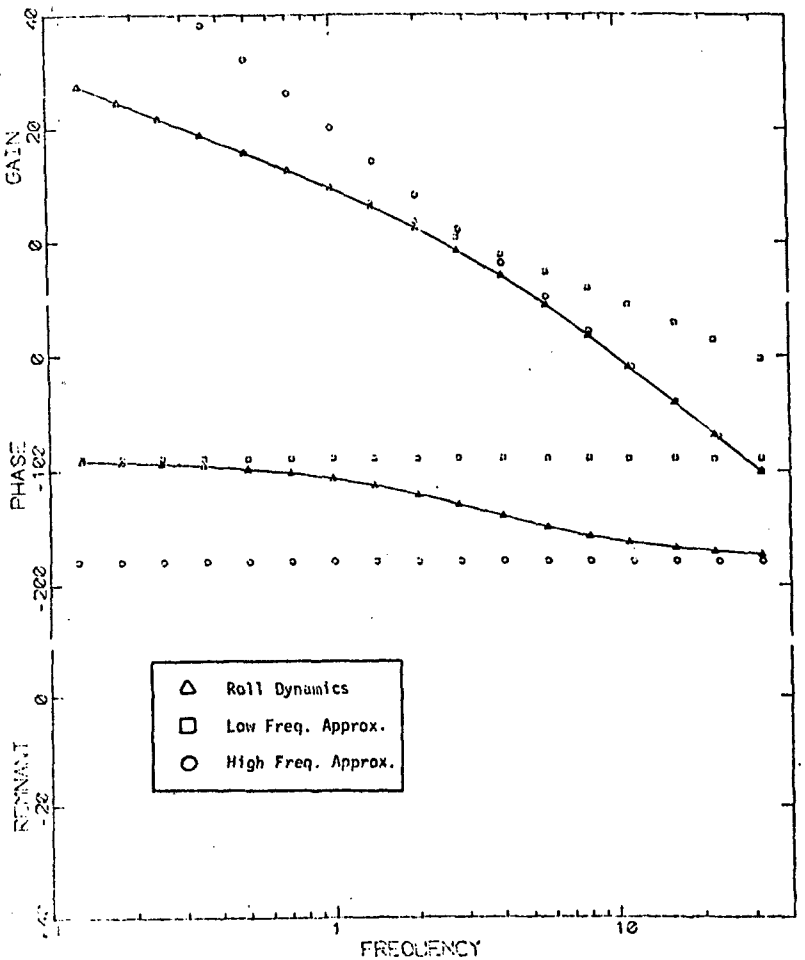


Figure 4. Vehicle Dynamics Approximations for High Performance Roll Dynamics

Learning the Input Characteristics

The K/s example described above was also examined to see if the effects of incorrect knowledge of the input characteristics would be evident. Figure 5 gives results for the case in which the operator overestimates the input bandwidth by a factor of 2. It can be seen that these results differ significantly from the situation in which the input bandwidth is known correctly only in terms of low frequency gain. If we refer to Figre 2, we see that the effect is also different from that of having an incorrect model of the vehicle dynamics.

Perceptual Efficiency

A major application of the wrong internal model concept would be to the study of learning of control strategies. In addition to learning the plant dynamics, it is believed that skill development involves learning to use the available cues most efficiently. We can envision this as a process of proper cue selection as well as noise reduction. For example, the progression-regression hypothesis [3] suggests an increasing utilization of derivative information with learning. It is therefore of interest to compare the effects of inefficient cue utilization and an incorrect internal model. Figure 6 compares predicted describing functions and remnant for optimized performance with and without rate information. The results are for the roll dynamics described earlier and it is assumed that the operator has learned the plant dynamics. It can be seen that failure to utilize rate information has a distinct impact on the measures of control strategy and perceptual efficiency. Most of this impact is at high frequencies, as expected. Furthermore, comparison with Figure 3 reveals that lack of rate information produces a decidedly different result from that obtained with a low frequency approximation to the vehicle dynamics. Thus, it should be

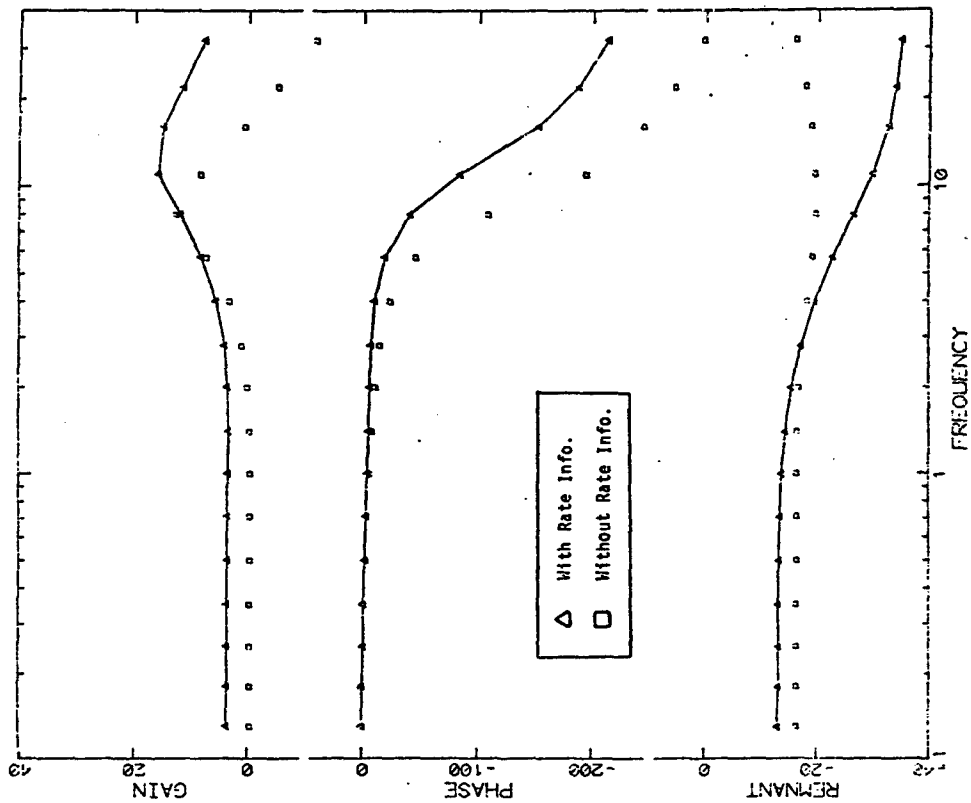


Figure 6. Effect of Non-Utilization of Rate Information on Roll Regulation

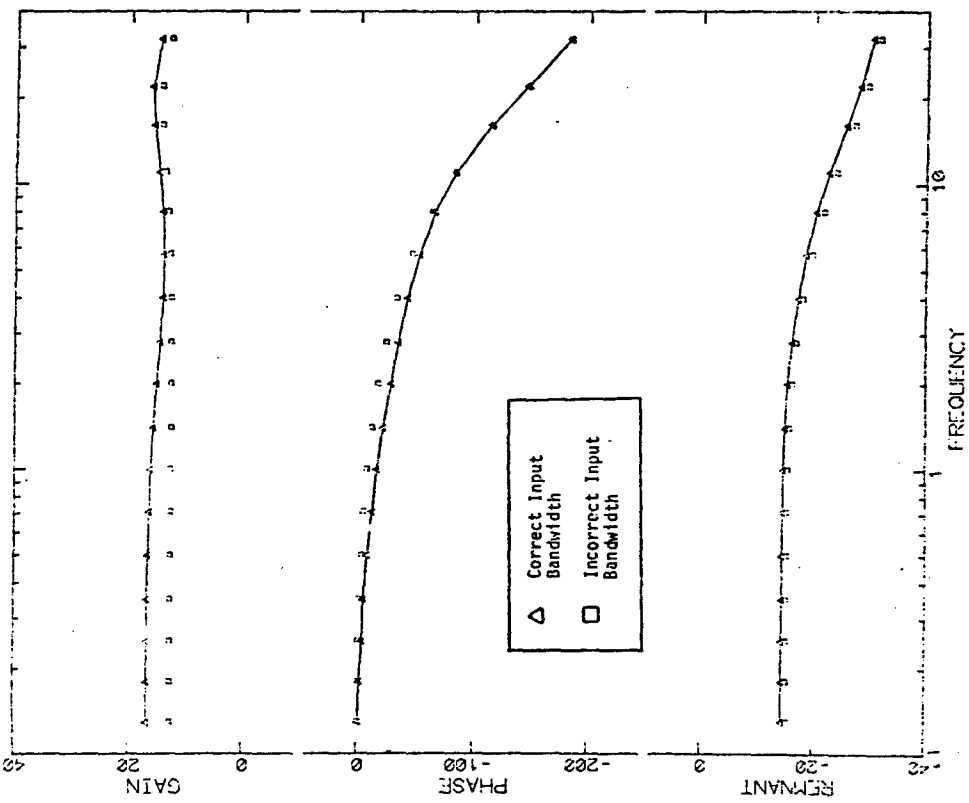


Figure 5. Effect of Wrong Estimate of Input Bandwidth on K/s Regulation

possible to differentiate between learning vehicle dynamics and learning to use the available cues from these measures of operator performance.

4. Conclusion

In conclusion, we wish to point out that, while the notion of a deviate internal model is appealing intuitively, in the authors' opinion, its use for trained operators even in complex tasks should be considered with caution for the following reasons: (1) the assumption of a perfect model works quite well for trained operators performing well defined tasks; (2) the observation and motor noises included in the optimal control model already account for some model imperfections; (3) when there is significant process noise, state prediction and estimation is difficult and the contribution to performance degradation of deviate internal models is likely to be reduced significantly; (4) computational requirements for predicting closed loop performance may well increase under this assumption; and (5) most importantly, in order to avoid having to choose among an infinity of possible internal models, rules for picking a specific internal model are needed, and, presently, no such rules exist. On the other hand, the programs developed here and the insights provided by the sensitivity analyses should prove very useful in studying and analyzing the performance of untrained operators.

5. References

1. Kleinman, D.L., S. Baron and W.H. Levison "A Control Theoretic Approach to Manned-Vehicle Systems Analysis" IEEE Trans. on Auto. Control, Vol. AC-16, No. 6, December 1971.

2. Baron, S. and J. Berliner, "MANMOD 1975: Human Internal Models and Scene-Perception Models," U.S. Army Missile Command, Redstone Arsenal, Alabama, Technical Report No. RD-CR-76-3, September 1975.

3. Fuchs, A.H. "The Progression-Regression Hypotheses in Perceptual-Motor Skill Learning" Journal of Experimental Psychology, 63, 1962, pp 177-182.

N79-17478

THEORETIC ASPECTS OF THE IDENTIFICATION OF THE PARAMETERS IN THE OPTIMAL CONTROL MODEL.

Ron A. van Wijk [*], Jan J. Kok

Man-Machine Systems Group, Laboratory for Measurement and Control, Department of Mechanical Engineering, Delft University of Technology, Delft, Netherlands.

0. Summary.

The identification of the parameters of the optimal control model (Baron, Kleinman, Levison [1]) from input-output data of the human operator is considered. Accepting the basic structure of the model as a cascade of a full-order observer and a feedback law, and suppressing the inherent optimality of the human controller, the parameters to be identified are the feedback matrix, the observer gain matrix and the intensity matrices of the observation noise and the motor noise.

The identification of the parameters is a statistical problem, because of the fact that the system and output are corrupted by noise, and therefore the solution must be based on the statistics (probability density function) of the input and output data of the human operator. However, based on the statistics of the input-output data of the human operator, no distinction can be made between the observation and the motor noise, which shows that the model suffers from overparametrization. In order to obtain a unique set of parameters, for the model to be identified an equivalent system must be defined, the associated system, in which the observation noise and the motor noise are replaced by an innovations process, which is a combination of these noises.

The parameters in the associated system can be identified if the following conditions are satisfied. First, the input and output signal must be persistently exciting of sufficient order. Second, the parameters of the associated system must satisfy a controllability and an observability condition. This last condition is plausible, for the non-controllable and non-observable part of the system do not influence the output of the system. As a consequence, in an identification procedure only the controllable and observable part of the system can be identified. But even if all the conditions are satisfied, the parameters of the associated system are not uniquely related to the model parameters. In fact the model parameters can be derived from the parameters of the associated system up to a similarity transformation of the model. The optimal control model exhibits a particular structure, due to the cascade of an observer and a feedback law. So,not every similarity transformation on the parameters of the associated system yields a set of parameters corresponding to this structure of the optimal control model. In order to satisfy this requirement the transformation matrix must be a solution of a certain quadratic matrix equation.

Once the parameters are determined, the optimality of the estimated parameters is considered making use of the results of the iverse optimal control problem and its filtering equivalent. The optimality conditions may further restrict the equivalent class of feedback matrices and observer gain matrices. Corresponding to each feedback law a class of weighting matrices exists, just as with the observer gain matrix and the corresponding noise intensity matrices. So the weighting matrices and the observation noise and motor noise cannot be identified uniquely without additional information about these parameters.

[*]) The research reported in this paper is partially supported by the Netherlands Organization for Advancement of Pure Research (ZWO).

1. Introduction.

A very important feature of the optimal control model is the fact that the parameter values which give the best description of the human controller, do satisfy certain general rules, which means that the model in conjunction with these rules for adjustment of its parameters can be used to predict the human controller behavior in a broad class of control situations.

As is known from publications of Kleinman, Baron, Levison [1] successful methods exist for iterative adjustment of the model parameters such that the model agrees adequately with human control behavior. However, this method for the determination of the model parameters also shows some aspects which can be prohibitive for application of the model in investigation of certain questions related to human control behavior. For example, to determine the optimal control law in the model it is assumed that the control task, as performed by the human controller, is known. For those situations where the appointed task is known in advance, and where also care has taken that the human controller optimizes this task,and no other, the given method can be used indeed. However, if one tries to examine the subjective interpretation of the human controller of a certain control task laid upon him, then the optimized performance index is *not* known a priori. In this case one would like to take the reverse approach by determining the control law carried on by the operator from the measurements of his input and output, then verifying the optimality of the control law, and subsequently establishing the corresponding performance index which has been optimized.

Not only in the control part, but also in the observation part of the model such difficulties can arise. According to the rules of parameter adjustment of Baron et. al., the intensity of the observation noise has a certain fixed connection with the power density of the observer signal. This statement is based on the assumption that the display used for observation of the system output is a standard design which fulfills standard norms. However, if one considers the question of the influence of display parameters on the corresponding observation noise, then again one is confronted with the reverse problem of determining the intensity of the observation noise directly from the measurements. Additional to this problem at the observation side,the motor side shows a similar problem related to the controls and the motor noise.

The problems mentioned above have motivated us to concentrate the investigation of the optimal control model on the development of an identification procedure, which allows for a direct estimation of the parameters of the model from measurements of the input and output of the human controller.

It should be mentioned that with respect to the general optimal control model as introduced by Baron, Kleinman and Levison,a simplification will be

introduced in the form of the suppression of the inherent time delay of the human controller and, consequently, of the predictor in the model. Such a simplification might be inadmissible for the application of the model in relatively fast vehicle control situations (like automobile and aircraft control), where the inherent time delay plays an important role, but seems to be acceptable for application of the model in relatively slow control situations; particularly the application of this simplified model in supervisory control.

2. *Structure of the model and the modal equations.*

Referring to Fig 1 which gives the structure of the model, the closed loop system consisting of the controlled system and the human operator model, is described by the following set of equations:

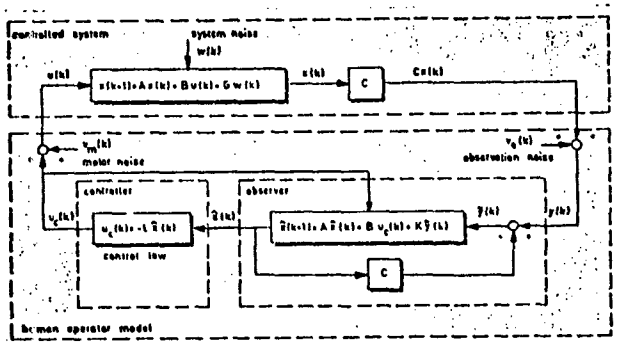


Fig. 1: Structure of the human operator model.

- State equation of the controlled system

$$x(k+1) = Ax(k)+Bu(k)+Gw(k) \; ; \tag{2.1}$$

- observation equation

$$y(k) = Cx(k)+v_o(k) \; ; \tag{2.2}$$

- state equation of the observer

$$\hat{x}(k+1) = A\hat{x}(k)+Bu_c(k)+K(y(k)-C\hat{x}(k)) \; ; \tag{2.3}$$

- control law

$$u_c(k) = -L\hat{x}(k) \; ; \tag{2.4}$$

- control equation

$$u(k) = u_c(k)+v_m(k) \; . \tag{2.5}$$

The various quantities involved are defined as follows:

$x(k)$ = state of the controlled system (n- dimensional);
$u(k)$ = input of the controlled system (m- dimensional);
$w(k)$ = system noise (stationary d_w- dimensional white noise with intensity matrix ψ_w);
$v_o(k)$ = observation noise (stationary d_y- dimensional white noise with intensity matrix ψ_{v_o});
$y(k)$ = observed output variable (d_y- dimensional);
$\hat{x}(k)$ = state of the observer (the one step predictor estimate $\hat{x}(k/k-1)$);
$u_c(k)$ = commanded input;
$v_m(k)$ = motor noise (stationary m- dimensional white noise with intensity matrix ψ_{v_m});
A, B, C, G = a priori known system matrices;
L = feedback matrix;
K = observer gain matrix.

The noises $w(k)$, $v_o(k)$ and $v_m(k)$ are assumed to be mutually uncorrelated.

If L is equal to the optimal feedback matrix L^o which optimizes the performance index:

$$J = \lim_{K\to\infty} \frac{1}{K} E\{ \sum_{k=k_o}^{K} x^T(k)Qx(k)+u^T(k)Ru(k)\} \; , \tag{2.6}$$

then:

$$L^o = R^{-1}B^TA^{-T}(P-Q) = (R+B^TPB)^{-1}B^TPA \; , \tag{2.7}$$

where the matrix P is the unique symmetric positive definite solution of the algebraic Riccati equation:

$$P = Q+A^TP[I+BR^{-1}B^TP]^{-1}A = Q+A^TPA-A^TPB[B^TPB+R]^{-1}B^TPA \; . \tag{2.8}$$

The weighting matrices Q and R in the performance index (2.6) are assumed to be symmetric and (semi) positive definite.

If K is equal to the optimal observer gain matrix K^o which minimizes the mean quadratic reconstruction error (the matrix W is a symmetric positive definite matrix):

$$J = E\{\tilde{x}^T(k)W\tilde{x}(k)\} \; , \tag{2.9}$$

where $\tilde{x}(k)$ is the reconstruction error:

$$\tilde{x}(k) = x(k)-\hat{x}(k) \; , \tag{2.10}$$

then:

$$K^o = AV_{\tilde{x}}C^T[\psi_{v_o}+CV_{\tilde{x}}C^T]^{-1} = [V_{\tilde{x}}-G\psi_wG^T-B\psi_{v_m}B^T]A^{-T}C^T\psi_{v_o}^{-1} \; , \tag{2.11}$$

where the matrix $V_{\tilde{x}}$, the covariance matrix of the reconstruction error, is the unique symmetric positive definite solution of the algebraic Riccati equation:

$$V_{\tilde{x}} = B\psi_{v_m}B^T+G\psi_wG^T+AV_{\tilde{x}}[I+C^T\psi_{v_o}^{-1}CV_{\tilde{x}}]^{-1}A^T =$$

$$B\psi_{v_m}B^T+G\psi_wG^T+AV_{\tilde{x}}A^T-AV_{\tilde{x}}C^T[CV_{\tilde{x}}C^T+\psi_{v_o}]^{-1}CV_{\tilde{x}}A^T \; . \tag{2.12}$$

The intensity matrices ψ_w, ψ_{v_m}, ψ_{v_o} are assumed to be symmetric and (semi) positive definite.

In the case that the performance index also includes the control rate the model remains of the form (2.3) - (2.5) by proper augmentation of the state vector.

3. *The choice of model parameters in relation to the identifiability.*

Having fixed the overall structure of the model, the parameters of the model will be considered now. Following the line of the optimal control model of Baron, Kleinman and Levison the parameter set consists of the set of matrices $\{Q,R,\psi_{v_m},\psi_{v_o}\}$, so the parameters of the performance index and of the statistics of the noises. In that case the feedback matrix L^o and the observer gain matrix K^o are considered as dependent parameters (L^o is the optimal feedback matrix corresponding to the weighting matrices Q and R, and K^o is the optimal observer gain matrix corresponding to the intensity matrices ψ_{v_m} and ψ_{v_o}). However, if the optimality of the human controller is not a priori accepted, then the parameter set consists of the matrices $\{L,K,\psi_{v_m},\psi_{v_o}\}$. The questions whether a given feedback matrix L is indeed optimal in the sense of some quadratic performance index and, if so, what is the corresponding set of weighting matrices Q and R, are then considered separately. The same counts for the optimality of a given observer gain matrix K and the corresponding intensity matrices ψ_{v_m} and ψ_{v_o}. This way of approach has some important advantages in the identification of the model from input-output data, like (possibly) a better fit of the data (a broader class of matrices L and K is allowed) and less problems with the uniqueness of the solution (a particular matrix L or K corresponds with a whole class of possible matrices Q and R or matrices ψ_{v_m} and ψ_{v_o}. The test whether the matrices L and K are solutions of an optimal control problem or observer problem respectively, can be verified by application of the results of the inverse optimal control problem and observer problem respectively (Sec. 9). This approach means basically that the structure of the optimal control model is fully accepted and adopted, however, that the optimality of the human controller is not assimilated in advance, and that it will be examined afterwards.

Based on the considerations given above the identification problem of the human operator model is defined in terms of the parameter set $\{L,K,\psi_{v_m},\psi_{v_o}\}$. In the evaluation of the identifiability of this parameter set it is assumed that there are experimental measurements available of the closed loop system (Fig 1), particularly of the control signal u(k) and the output signal Cx(k).

The system equation and the output equation of the system to be identified (2.3) - (2.5) can be written in the following form:

$$\hat{x}(k+1) = (A-KC-BL)\hat{x}(k)+KCx(k)+Kv_o(k) \; ; \tag{3.1}$$

$$u(k) = -L\hat{x}(k)+v_m(k) \; . \tag{3.2}$$

In the system (3.1), (3.2), the input Cx(k) and the output u(k) are measured and the system matrices A,B,C and G are assumed to be known.

In the following sections the problem will be treated whether the parameters L,K,ψ_{v_m} and ψ_{v_o} of the system (3.1), (3.2) operating in a closed loop are identifiable and under what conditions. First some general results of the identifiability concept will be presented, which then will be applied to the underlying problem.

4. *General results of the identifiability concept.*

Before solving the parameter identification problem, the problem of the identifiability of the parameters should be considered. In case of the model structure is given, the question also arises whether the input signals are appropriate for the identification of the unknown parameters in the system. The identifiability problem can be approached in a deterministic way or a stochastic way. Tse and Anton [2] argue that deterministic identifiability is necessary but not sufficient for stochastic identifiability when the output and (or) the system is corrupted by noise. This is plausible, for in a deterministic way nothing can be said about the statistics of the noise sources. So, in the given stochastic case the system parameters have to be recovered in some probabilistic sense.

For the elaboration of the identifiability we introduce the following notation:
The observations u(k) will be denoted by z(k); the set of observations z(l) for $k_o \le l \le k$ will be written as Z(k):

$$Z(k) \triangleq \{z(k_o), \ldots, z(k)\} \; ; \tag{4.1}$$

the vector of unknown parameters will be denoted by θ, in the underlying problem $\theta = \{K,L,\psi_{v_m},\psi_{v_o}\}$; (4.2)

the probabilitydensity function of z(k) conditioned on Z(k-1) and θ is: $p(z(k)/Z(k-1),\theta)$.

Based on statistical considerations, Tse and Anton [2] defined resolvability of two parameters in the following way:

Two parameters θ_1 and θ_2 belonging to some compact (s- dimensional) set R_c^s are *unresolvable* if the equality:

$$p(z(k)/Z(k-1),\theta_1) = p(z(k)/Z(k-1),\theta_2) \tag{4.3}$$

holds with probability 1 for all, except a finite number of integers $k \ge k_o$.

Under some mild conditions resolvability guarantees the existence of a consistent sequence of estimates, which means the identifiability of parameters in a probabilistic sense. One of these conditions is the requirement that the observation statistics z(k) are asymptotically uncorrelated. If the closed loop system is stable, which will be the case in a normal control situation, this condition is always satisfied. It easily can be derived that in that case A-BL and A-KC are stable matrices.

So, in this section the conclusion is drawn that the identifiability of the parameters have to be based on the probability density function of the measurements. Tse and Weinert [3] have applied this concept to a linear system without an additional input signal. In the following section the results will be applied to the problem of the identifiability of the parameters in the system given by (3.1) and (3.2), which is more difficult because of the extra input Cx(k).

5. *Application of the identifiability results to the linear system.*

The closed loop system (2.1) - (2.5) is linear and the noise sources are assumed to be Gaussian. In this case the probability density function $p(z(k)/Z(k-1),\theta)$ is completely characterized by its mean and variance:

$$\hat{z}(k/k-1,\theta) \triangleq E\{z(k)/Z(k-1),\theta\} \; ; \tag{5.1}$$

$$V_z(k/k-1,\theta) \triangleq E\{(z(k)-\hat{z}(k/k-1,\theta))(z(k)-\hat{z}(k/k-1,\theta))^T/Z(k-1),\theta\}. \tag{5.2}$$

In the stationary control situation the variance has a constant value which will be denoted by V_z. According to Eq.(4.3), the identifiability of the parameters now can be related to $\hat{z}(k/k-1,0)$, and to V_z. So, two parameters θ_1 and θ_2 are resolvable if:

$$\hat{z}(k/k-1,\theta_1) \neq \hat{z}(k/k-1,\theta_2) \; ; \tag{5.3}$$

$$V_z(\theta_1) \neq V_z(\theta_2) \; . \tag{5.4}$$

for all except a finite number of integers $k \geq k_o$.

For the linear system to be identified with Gaussian inputs :

$$\hat{x}(k) = A^*\hat{x}(k)+KCx(k)+Kv_o(k) \; ; \tag{3.1}$$

$$z(k) \triangleq u(k) = -L\hat{x}(k)+v_m(k) \; ; \tag{3.2}$$

$$A^* \triangleq A-KC-BL \; , \tag{5.5}$$

the mean (5.3) and variance (5.4) can be calculated. A Kalman filter (one stage predictor) of the system (3.1), (3.2) generates a linear projection of the state $\hat{x}(k)$ on the measurement space $Z(k-1)$ which for the linear Gaussian case is equal to:

$$\xi(k) \triangleq E[\hat{x}(k)/Z(k-1),\theta] \; . \tag{5.6}$$

So a possible way to generate $\hat{z}(k/k-1,\theta)$ is:

$$\hat{z}(k) \triangleq \hat{z}(k/k-1,\theta) = -L\xi(k) \; . \tag{5.7}$$

Here $\xi(k)$ is the output of the Kalman filter:

$$\xi(k+1) = A^*\xi(k)+KCx(k)+Fv(k) \; ; \tag{5.8}$$

$$z(k) = -L\xi(k)+v(k) \; . \tag{5.9}$$

In this algorithm the observed signal $Cx(k)$ represents a known input to the filter, $v(k)$ is the innovations process, which is a white noise process with intensity ψ_v, and the matrix F is the filter gain matrix. The variance V_z results from the Eqs. (5.2), (5.7) and (5.9):

$$V_z = \psi_v \; . \tag{5.10}$$

As mentioned before the algorithm (5.7) - (5.9) is just one possible way to obtain $\hat{z}(k)$. It is clear by direct substitution that any system related to (5.7) - (5.9) in the following way (for every non-singular T):

$$\hat{A}^* = TA^*T^{-1}; \; \hat{K} = TK; \; \hat{F} = TF; \; \hat{L} = LT^{-1} \tag{5.11}$$

generates the same $\hat{z}(k)$ and V_z. Thus,the system with which $\hat{z}(k)$ and V_z can be calculated will be written by:

$$\hat{\xi}(k+1) = \hat{A}^*\hat{\xi}(k)+\hat{K}Cx(k)+\hat{F}v(k) \; ; \tag{5.12}$$

$$z(k) = -\hat{L}\hat{\xi}(k)+v(k) \; ; \tag{5.13}$$

$$\hat{z}(k) = -\hat{L}\hat{\xi}(k) \; . \tag{5.14}$$

As a first result ,it can be noticed that in consequence of the statistical approach it is not possible, from measurements $Cx(k)$ and $z(k)$ only, to distinguish between the system (5.12), (5.13) and the system to be identified (3.1), (3.2). In the system (5.12), (5.13), which is called the *associated system*, the observation noise and the motor noise are replaced by the equivalent white noise source $v(k)$ and the gain matrix $\hat{F}$. So the conclusion can be drawn that, from the measurements $Cx(k)$ and $z(k)$, it is *not* possible to determine each of the intensities of the observation noise ψ_{v_o} and of the motor noise ψ_{v_m},seperately. Only the equivalent white noise source $v(k)$ and the gain matrix can be identified instead. This conclusion implies that the parameter set to be identified has to be adjusted to the parameters of the associated system:

$$\hat{\theta} = \{\hat{A}^*, \hat{K}, \hat{L}, \hat{F}, \hat{\psi}_v\} \; . \tag{5.15}$$

On behalf of the application of the identifiability results (5.3), (5.4) to the associated system, this system will be rewritten in the following form:

$$\xi(k+1) = (\hat{A}^*+\hat{F}\hat{L})\xi(k)+[\hat{K}|\hat{F}]\gamma(k) \; ; \tag{5.16}$$

$$\hat{z}(k) = -\hat{L}\xi(k) \; . \tag{5.17}$$

Here the input signal $\gamma(k)$ is a composition of the measured input and output of the system to be identified, and it is completely known:

$$\gamma(k) = \begin{bmatrix} Cx(k) \\ z(k) \end{bmatrix} \; . \tag{5.18}$$

Since the closed loop system is assumed to be stable, also $(\hat{A}^*+\hat{F}\hat{L})$ is a stable matrix, so for the stationary situation the response of the system (5.16), (5.17) can be written as:

$$\hat{z}(k) = \sum_{i=0}^{k-k_o-1} -\hat{L}(\hat{A}^*+\hat{F}\hat{L})^i[\hat{K}|\hat{F}]\gamma(k-i-1) \; , \tag{5.19}$$

or:

$$\hat{z}(k) = \sum_{i=0}^{k-k_o-1} H(i/\hat{\theta})\gamma(k-i-1) \; , \tag{5.20}$$

where theimpulse response matrix $H(i/\hat{\theta})$ is defined by:

$$H(i/\hat{\theta}) = -\hat{L}(\hat{A}^*+\hat{F}\hat{L})^i[\hat{K}|\hat{F}] \; . \tag{5.21}$$

From the identifiability condition (5.3) and Eq (5.20) it follows that if two different impulse response matrices $H(i/\theta_1)$ and $H(i/\theta_2),(i\geq 0)$, with the same input $\gamma(k)$ give the same response $\hat{z}(k)$ the parameters θ_1 and θ_2 are unresolvable. Thus from condition (5.3) the conclusion can be drawn that for a given input $\gamma(k)$ and given output $\hat{z}(k)$ the impulse response matrix $H(i/\hat{\theta})$, $i\geq 0$, has to be unique. In the next section this will be worked out in detail. Then also the question should be considered under what conditions the parameters , from the impulse response matrix, can be resolved uniquely. This problem will be solved in Sec.7.

From the identifiability condition (5.4), and the Eqs. (5.13) and (5.14), the uniqueness of ψ_v follows as a result.

6. *Uniqueness of the impulse response matrix.*

We will now examine under what conditions the impulseresponse matrix $H(i/\hat{\theta})$, $i\geq 0$, is unique for a given input $\gamma(k)$ and a unique output signal $\hat{z}(k)$, $k\geq k_o$. Since the parameters to be identified are determined by the impulse responsematrix (5.21), the side constraint has to be posed that

$H(i/\hat{\theta})$, $i>0$, is a realizable impulse response matrix of a finite dimensional time invariant system. In this case it follows from the Cayley-Hamilton theorem that a matrix $(\hat{A}^*+\hat{F}\hat{L})^i$, $i\geq n$, is a linear function of the matrices $(\hat{A}^*+\hat{F}\hat{L})^i$, $i=0,\ldots,n-1$. Thus, it is only necessary to work out the conditions for the set $H(i/\hat{\theta})$, $i=0,\ldots,n-1$, to be unique.

Suppose now that two impulse response matrices $H_1(i/\hat{\theta})$ and $H_2(i/\hat{\theta})$ satisfy the Eq.(5.20). Subtraction of both the equations leads to the result:

$$0 = \sum_{i=0}^{k-k_o-1} (H_1(i/\hat{\theta})-H_2(i/\hat{\theta}))\gamma(k-i-1) \quad . \tag{6.1}$$

From Eq (6.1) it can be concluded that:

$$[\Delta H(0)\ldots\Delta H(k-k_o-1)]\begin{bmatrix} E\{\gamma(k-1)\gamma^T(k-1)\}\ldots E\{\gamma(k-1)\gamma^T(k_o)\} \\ \vdots \\ E\{\gamma(k_o)\gamma^T(k-1)\}\ldots E\{\gamma(k_o)\gamma^T(k_o)\} \end{bmatrix}\begin{bmatrix} \Delta H^T(0) \\ \vdots \\ \Delta H^T(k-k_o-1) \end{bmatrix} = 0 \; ; \tag{6.2}$$

with:

$$\Delta H(\;) = H_1(j/\hat{\theta})-H_2(j/\hat{\theta}) \quad . \tag{6.3}$$

Because of the fact that only the uniqueness of $H(i/\hat{\theta})$, $i=0,\ldots,n-1$ must be shown, $k = k_o+n$ will be choosen. Then, if:

$$\begin{bmatrix} E\{\gamma(k_o+n-1)\gamma^T(k_o+n-1)\}\ldots E\{\gamma(k_o+n-1)\gamma^T(k_o)\} \\ \vdots \\ E\{\gamma(k_o)\gamma^T(k_o+n-1)\} \quad \ldots E\{\gamma(k_o)\gamma^T(k_o)\} \end{bmatrix} > 0 \; ; \tag{6.4}$$

(>0 means: the matrix is positive definite) from Eq (6.2) it follows that:

$$[\Delta H(0)\ldots\Delta H(k-k_o-1)] = 0 \quad , \quad k = k_o+n \quad . \tag{6.5}$$

From Eq. (6.5), by using Eq.(6.3) and the fact that $H(i/\hat{\theta})$, $i\geq n$, is a linear function of $H(i/\hat{\theta})$, $i=0,\ldots,n-1$, it results that the impulse response matrix $H(i/\hat{\theta})$ is unique. A signal $\gamma(k)$ with the property given by the Eq. (6.4) is called *persistently exciting of order (n-1)*. A sufficient condition for this property can also be given in the (discrete) frequency domain. The signal $\gamma(k)$ is persistently exciting of infinite order if:

$$S_{\gamma\gamma}(z) > 0, \quad z = e^{j\omega}; \quad 0 \leq \omega \leq 2\pi \quad , \tag{6.6}$$

where $S_{\gamma\gamma}(z)$ is the discrete spectral density matrix of $\gamma(k)$. From the condition (6.6) it is easily seen that if $\gamma(k)$ is a white noise source, it is persistently exciting of infinite order. This was the case Tse and Weinert [3] have considered. In our case it is not possible, only based on the structure of the system to be identified, to show directly that $\gamma(k)$ is persistently exciting of sufficient order. But, in the given situation the signals $z(k)$ and $Cx(k)$, and thus $\gamma(k)$, are known, so it is possible to test the condition (6.4). If this condition is satisfied, the impulse response matrix $H(i/\hat{\theta})$, $i=0,1,\ldots$ turns out to be unique. However, the conditions (6.4) and (6.6) are only sufficient, so if (6.4) (or (6.6)) is not satisfied, it does not imply that the impulse response matrix $H(i/\hat{\theta})$ is not unique.

7. *Realization of the impulse response matrix.*

After application of the test (6.4) for the measured signals the next step in the identifiability problem is to work out the uniqueness and the structure of the parameters given the uniqueness of the impulse response matrix. This problem is known in literature as the *realization problem*. An extensive treatment of this problem is given by Silverman [4].

The following definitions will be used:

Definition 7.1:
A system $(A,B,C)_n$ *) is *H-equivalent* to $(\bar{A},\bar{B},\bar{C})_n$ if and only if the impulse response matrices of both the systems are equal.

Definition 7.2:
A system $(A,B,C)_n$ is *strictly equivalent* to $(\bar{A},\bar{B},\bar{C})_n$ if and only if there exists a constant non-singular matrix T such that:

$$\bar{A} = TAT^{-1}; \; \bar{B} = TB; \; \bar{C} = CT^{-1} \quad . \tag{7.1}$$

Definition 7.3:
The system $(A,B,C)_n$ with dimension n is called *minimal* if no H-equivalent system $(\bar{A},\bar{B},\bar{C})_{\bar{n}}$ exists with dimension $\bar{n} < n$.

Using the concepts given above, Silverman proved the following theorem.

Theorem 7.1:
Minimal time-invariant system representations are strictly equivalent if and only if they are H-equivalent.

Applied to the situation in which the impulse response matrix $H(i/\hat{\theta})$ is given by Eq.(5.21), this theorem gives the result that *every* minimal H-equivalent system of $(\hat{A}^*+\hat{F}\hat{L},[\hat{K}|\hat{F}],-\hat{L})_n$ is given by:

$$\hat{\bar{A}}^*+\hat{\bar{F}}\hat{\bar{L}} = T(\hat{A}^*+\hat{F}\hat{L})T^{-1}; \; [\hat{\bar{K}}|\hat{\bar{F}}] = T[\hat{K}|\hat{F}]; \; -\hat{\bar{L}} = -\hat{L}T^{-1} \quad . \tag{7.2}$$

In Sec.5 (eqs.(5.8) and (5.9)) it is argued that the system $(A^*+FL,[K|F],-L)_n$ is one of the strictly equivalent systems, so statement (7.2) can also be put in the following form (after proper adjustment of the transformation matrix T):

Every minimal H-equivalent system of $(A^*+FL,[K|F],-L)_n$ is given by:

$$\hat{\bar{A}}^*+\hat{\bar{F}}\hat{\bar{L}} = T(A^*+FL)T^{-1}; [\hat{\bar{K}}|\hat{\bar{F}}] = T[K|F]; \; -\hat{\bar{L}} = -LT^{-1}. \tag{7.3}$$

The above statement is equivalent with:

$$\hat{\bar{A}}^* = TA^*T^{-1}; \; \hat{\bar{K}} = TK; \; \hat{\bar{F}} = TF; \; \hat{\bar{L}} = LT^{-1}. \tag{7.4}$$

From the total class of possible parameters (7.4) a unique set must be selected. This is considered in the following theorem.

Theorem 7.2:
The matrices $\hat{A}^*,\hat{K},\hat{F},\hat{L}$ can be uniquely identified from the unique impulse response matrix $H(i/\hat{\theta})$ if and only if:

$$\text{o } (\hat{A}^*,L) \text{ is a reconstructable pair,} \tag{7.5}$$

*) The notation $(A,B,C)_n$ represents a linear time invariant n-dimensional system with system matrix A, input matrix B and output matrix C.

- $(\hat{A}^*,[K|F])$ is a controllable pair, (7.6)
- A canonical form for the system has been chosen. (7.7)

No strict proof of this theorem will be given. The conditions are made plausible in the following way:
If a matrix pair $(\hat{A}^*,\hat{L})$ is not reconstructable, then, according to Popov [5], the pair $(\hat{A}^*+\hat{F}\hat{L},\hat{L})$ is also not reconstructable. So, the not-reconstructable subsystem does not influence the output and, thus is not represented in the impulse response matrix. As a consequence, the non-reconstructable subsystem cannot be identified from the impulse response matrix. A similar argument holds for the controllability condition.
All minimal H-equivalent systems are given by Eq.(7.4), where T is some constant non-singular matrix. One way to guarantee the uniqueness of the parameter set is to choose a transformation matrix T such that a canonical parameter set results. For a proper canonical form only a unique transformation matrix T exists. The structure of the canonical form is determined by the so called *Kronecker invariants* (i.e. Popov [6]). The Kronecker invariants are given by the observability matrix of the system and remain unchanged under a similarity transformation. The consequence for the identification procedure is that prior to a parameter identification a *structural identification* must be carried out.

So far, no use has been made of the fact that the parameters of the controlled system (the matrices A, B, C and G) are given in advance, and thus that there exists a certain relation with the parameters to be identified. In particular:

$$\hat{A}^* = \widehat{(A-KC-BL)} = A-\hat{K}C-B\hat{L} \quad . \tag{7.8}$$

Then, by Eq.(7.4), the following quadratic equation results:

$$T\hat{B}\hat{L}T+\hat{A}^*T-TA+\hat{K}C = 0 \quad . \tag{7.9}$$

So, to obtain a set of parameters which satisfies the given structure of the model, the transformation matrix T must be an element of the finite set of solutions of the quadratic equation (7.9). Moreover, from observability considerations, it can be shown that this transformation matrix must be real. In Potter [7] the solution of a more specific quadratic equation is given, which can be generalized to the underlying equation. The maximum number of real regular solutions T which satisfy Eq.(7.9) is equal to $\binom{2n}{n}$. Using the controllability condition (7.5), it can be shown that to each real regular solution T one and only one triple $\hat{K}$, $\hat{F}$ and $\hat{L}$ corresponds. Thus, there are as many solutions of the realization problem as there are solutions T of Eq.(7.9), but in general *no* system in canonical form will belong to the set of possible solutions. A way to work out the identification is to identify the structure of the system, then to choose a canonical form,, to identify the parameters in the accepted canonical form, and to transform the canonical parameters using Eq.(7.9) to the ultimate parameter set. It should be noticed that the set of solutions is not influenced by the particular choice of the canonical form (with proper invariants).

In the introduction (Sec.1) it was argued that, once the parameters K and L are identified, the optimality of these parameters should be verified which might restrict the total set of possible solutions. In the following section some results will be presented in relation to this problem.

8. *The optimality of the feedback matrix and the observer gain matrix.*

So far, no assumption was made relative to the optimality of the feedback matrix L and the observer gain matrix K. In this section some results of the so called inverse optimal control and observer problem, which is considered in more detail in van Wijk, Kok [9], will be applied to the identified parameters L and K. The main part of the inverse optimal control problem concerns the question under which conditions a given feedback matrix L, in relation to the system matrices A and B, minimizes a quadratic cost function (2.6).
A similar question holds for the inverse observer problem, i.e. under which conditions a given gain matrix K, in relation to the matrices A and C, minimizes the quadratic reconstruction error (2.9).

For the control part it can be shown that if the matrix:

$$LA^{-1}B \text{ is } \textit{symmetrizable} \quad , \tag{9.1}$$

the feedback matrix L minimizes some quadratic criterion of the form (2.6). The concept of symmetrizability is defined by Taussky [10] in the following way:

> A matrix is symmetrizable if it is similar to a symmetric matrix, or, equivalently, if it has real characteristic roots and a full set of characteristic vectors.

The symmetrizability requirement is plausible indeed, for an optimal feedback matrix L^o does satisfy Eq.(2.7), so:

$$RL^o = B^TA^{-T}(P-Q) \quad . \tag{9.2}$$

With the symmetry of P and Q it follows that $RL^oA^{-1}B$ is a symmetric matrix. From the fact that R is positive definite it follows that $L^oA^{-1}B$ is similar to a symmetric matrix, hence the matrix is symmetrizable.

If the symmetrizability condition is satisfied, only the positive definiteness of the weighting R of the input is guaranteed, but the weighting matrix Q of the state is not necessarily semi positive definite. If, in addition, Q is required to be semi positive definite, also the following two necessary conditions must be satisfied by L:

- The *rank condition:*

$$r(L) = r(LA^{-1}B) \; ; \tag{9.3}$$

- The *eigenvalue condition:*

$$0 \leq \lambda_i(LA^{-1}B) < 1 \quad . \tag{9.4}$$

The rank condition follows from the fact that the Riccati matrix P is also a semi positive definite matrix, and from the following relation, derived from the Eqs.(2.7) and (2.8):

$$(R+B^TPB)L^oA^{-1}B = B^TPB \quad . \tag{9.5}$$

The eigenvalue condition can be shown along the semi positive definiteness of P and another relation resulting from the Eqs.(2.7) and (2.8):

$$R(I-L^oA^{-1}B)^{-1}L^oA^{-1}B = B^TPB \geq 0 \quad . \tag{9.6}$$

For a given feedback matrix which is obtained by identification, the above conditions can be tested in order to examine its optimality. In [9] the complete class of weighting matrices Q and R belonging to L^o is deduced. From this, the conclusion can be drawn that without additional structural information about the weighting factors no unique pair R and Q can result.

Based on the duality between the inverse optimal control and observer problem the following conditions can be concluded for the optimality of the observer gain matrix K:

- The *symmetrizability condition:*

  The matrix $CA^{-1}K$ is symmetrizable . (9.7)

- The *rank condition:*

  $r(K) = r(CA^{-1}K)$. (9.8)

- The *eigenvalue condition:*

  $0 \leq \lambda_i\{CA^{-1}K\} < 1$. (9.9)

Once a matrix K is identified, the conditions given above can be used to check the optimality of the observer.

In Sec.8 it was concluded that no unique pair {L,K} can be the solution of the identification problem, but it was derived that:

$$\hat{L} = LT^{-1}, \quad \hat{K} = TK \tag{9.10}$$

The transformation matrix T is a real regular solution of the quadratic equation (7.9). The given optimality conditions must be checked for every pair $(\hat{L},\hat{K})$. Because of the fact that the optimality conditions are *not* invariant under the transformation (9.10), these conditions may restrict the total class of possible solutions. However, to each pair {L,K} for which the optimality test is successful, a whole class of weighting matrices in the performance index (2.6) corresponds. The same is true for a whole class of intensity matrices of the observation noise and motor noise, respectively. So, without additional information, no unique model parameter values Q, R, ψ_{v_o}, ψ_{v_m} can be identified from the input-output data of the human operator.

9. Discussion of the identification and optimality conditions.

The identifiability of the parameters was based on the probability density function of the measurements. In consequence of the statistical considerations, no distinction could be made between the system to be identified, Eqs.(3.1) and (3.2), and the associated system, Eqs.(5.12) and (5.13). It was argued that it is not possible to identify the observation noise and motor noise as independent noise sources, but instead, the intensity of an equivalent white noise innovations process v(k) and its input matrix F can be determined only. For the identification of the parameters a procedure can be used as described in Kok, van Wijk [8]. The parameter ψ_v can always be identified uniquely; the other parameters of the associated system can be identified under the following (sufficient) conditions:

a) The signal $\begin{bmatrix} Cx(k) \\ z(k) \end{bmatrix}$ is persistently exciting of order (n-1),

b) $(\hat{A}^*,[\hat{K}|\hat{F}])$ is a controllable pair,

c) $(\hat{A}^*,\hat{L})$ is a reconstructable pair .

d) A canonical form has been chosen for $(\hat{A}^*,[\hat{K}|\hat{F}],\hat{L})$.

Condition a) can be tested prior to the identification. However, if this condition is not satisfied it does not necessarily mean that the parameters cannot be identified at all.
It is not possible to test conditions b) and c) in advance. However, the controllability and the reconstructability must be considered as inherent properties of the results of an identification procedure.
In relation to condition d) it was argued that in general the parameters of the model to be identified will *not* represent a canonical structure. However, in the application of an identification procedure it still is possible to make use of an a priori determined canonical form with proper Kronecker invariants. Once the parameters of the model in canonical form are identified, they must be transferred to a set of parameters which fits the subsumed model structure. The quadratic equation (7.9) determines the set of possible (real regular) transformation matrices T which will provide for this conversion. In general more than one (maximal $\binom{n}{m}$) real regular solution T will exist. However, if the quadratic equation has *not* at least one real regular solution T, it turns out that *no* set of parameters with the subsumed structure in the model fits the input-ouput data. In this case the assumptions of the model structure must be abandoned.

In the above, the structure of the system to be identified (i.e. the Kronecker invariants) is assumed to be known in order to be able to identify the parameters. So, prior to the parameter identification, a structural identification must be worked out. In literature however, very few procedures are available for this purpose.

In conclusion it can be stated that in general more than one set of the parameters $\{K,L,\psi_v,F\}$ is the possible solution of the identification problem, and therefore some additional information should be available in order to determine a unique parameter set. Once the parameters L and K are determined, the optimality of these parameters can be checked using conditions (9.1), (9.3), (9.4) and (9.7) - (9.9), respectively.

As an overall conclusion it can be stated that, from the identification point of view, the optimal control model suffers from overparametrization. The observation noise and motor noise are not resolvable, and the feedback matrix and the observer gain matrix can only be determined up to a similarity transformation of the model, where the transformation matrix is restricted to the solution of a quadratic equation.

10. References.

1 Kleinman, D.L.; Baron, S.; Levison, W.H. A Control Theoretic Approach to Manned-Vehicle System Analysis. IEEE Trans. on A.C. Vol. AC-16 No. 6 (1971) pp. 824-832.

2 Tse, E.; Anton, J. On the Identifiability of Parameters. IEEE Trans. on A.C. Vol. AC-17 No. 5 (1972) pp. 637-646.

3 Tse, E.; Weinert, H. Correction and Extension of "On the Identifiability of Parameters". IEEE Trans. on A.C. Vol. AC-18 No. 6 (1973) pp. 687-688.

4 Silverman, L.M. Realization of Linear Dynamical Systems. IEEE Trans. on A.C. Vol AC-16 No. 6 (1971) pp. 554-567.

5 Popov, V.M. Hyperstability of Control Systems. (book) Springer-Verlag, Berlin, (1973).
6 Popov, V.M. Invariant Description of Linear Time-Invariant Controllable Systems. SIAM. J. Control Vol. 10 No. 2 (1972) pp. 252-264.
7 Potter, J.E. Matrix Quadratic Solutions. J. SIAM. Appl. Math. Vol. 14 No. 3 (1966) pp. 496-501.
8 Kok, J.J.; van Wijk, R.A. Identification of the Optimal Control Model. In: H.G. Stassen. Progress Report January 1973 until July 1976 of the Man-Machine Systems Group. Report: Delft, Delft Univ. of Techn., Lab. for Meas. and Contr., 1977, pp. 105-125 WTHD 95.
9 van Wijk, R.A.; Kok, J.J. The Inverse Optimal Control Problem and Observer Problem. In: H.G. Stassen. Progress Report January 1973 until 1976 of the Man-Machine Systems Group. Report: Delft, Delft Univ. of Techn., Lab. for Meas. and Contr., 1977, pp. 126-140 WTHD 95.
10 Taussky, O. Positive-Definite Matrices and Their Role in the Study of Characteristic Roots of General Matrices. Advances in Mathematics, Vol. 2 No. 2 (1968) pp. 175-186.

A QUASI-LINEAR CONTROL THEORY ANALYSIS OF TIMESHARING SKILLS

Diane Damos
SUNY at Buffalo

Christopher Wickens
University of Illinois

Human involvement with complex systems often requires the operator to timeshare or perform several tasks concurrently. While it is apparent that performance under multiple- or dual-task conditions often benefits dramatically from practice, the precise sources of improvement are not clear. One possible source is further mastery of the single-task skills themselves. A second possible source is timesharing skills, which are hypothesized to contribute only to multiple-task performance and which do not develop under single-task conditions. To date, the identity, development, and generality of timesharing skills have not been clearly isolated experimentally. The present study examines performance with practice on two dual-task combinations -- dual-axis tracking and two discrete information processing tasks -- in an effort to identify the presence and development of specific timesharing skills, such as parallel information processing or rapid intertask switching. The generality of timesharing skills also is investigated by examining transfer of these skills between the two qualitatively different task combinations.

METHOD

Design

The experiment employed three groups. The dual-task transfer group, Group 1, received dual-task training on Days 1 and 2. Group 2, the single-task transfer group, received single-task training on Day 1 and dual-task training on Day 2. Group 3, the control group, received dual-task training on Day 2 only. Sixteen subjects were assigned at random to each group. Subjects in Groups 1 and 2 performed two discrete information processing tasks (short-term memory and classification) on Day 1. All subjects performed two tracking tasks on Day 2.

Task

Short-Term Memory (STM). Random digits between one and four were presented sequentially to the subject. The subject retained the most recently displayed digit in memory while responding to the preceding digit. For example, if the first stimulus was a "1" and the second a "3", the correct response to "3" was "1". The subject entered her response via a four-choice keyboard attached to the right side of the experimental chair. As soon as the response was made, the stimulus was erased and the next one presented. The dependent variable was the average interval between correct responses (CRI).

Classification (CL). Two randomly selected digits between five and eight were presented simultaneously to the subject. The digits varied on two dimensions: size and name. The subject determined the number of dimensions on which the stimuli were alike and pressed one of three keys

on a keyboard attached to the left side of the subject's chair. As soon as the subject made a response, the pair was erased and a new pair presented. The dependent variable was again the CRI. Under dual-task conditions the visual angle subtended by the STM and CL tasks was 1.09° by .31°.

Tracking (TR). Two one-dimensional compensatory tracking tasks each required the subject to keep a cursor centered in a horizontal track by appropriate left-right manipulations of a control stick. The two tasks were identical except that one task was controlled by the subject's right hand; the other, by her left hand. The disturbance input to each display was a random forcing function with an upper cutoff frequency of .32 Hz. The inputs to the two tasks were statistically independent. The transfer function of each system was $Y = K\ (.25/S + .75/S^2)$. Average absolute error was recorded after each trial while the position of the control stick and the error cursor was recorded every 120 msec for later offline analysis.

Figure 1 shows the display for the TR-TR combination. Also presented on the display were two performance bars similar to those described by Wickens and Gopher (1975). The two tracking error displays subtended a visual angle of 4.05° by .70°.

Procedure

Day 1 Training. On Day 1, Group 1 received predominately dual-task training while Group 2 received training that was as similar as possible to that of Group 1 except that the subject never performed the tasks simultaneously. Subjects in each group received a total of 46, one-minute trials which were grouped into six blocks. During Block 1, the subjects alternated performing each task alone for a total of five trials on each

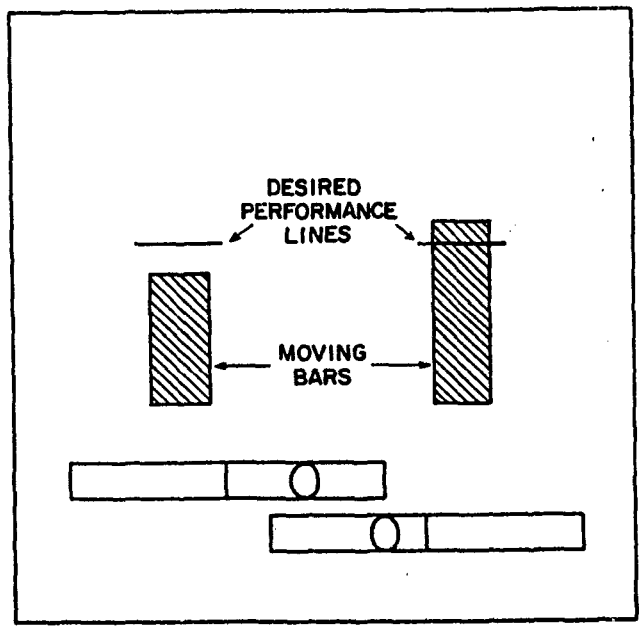


Figure 1. The dual-task tracking display.

task. Beginning with Block 2, the subjects in Group 1 received both single- and dual-task training. In each block, five dual-task trials were followed by one single-task trial on each task. The subjects in Group 2 continued to alternate between the two tasks throughout Day 1.

Day 2 training. Day 2 training was conducted on the day immediately following Day 1 training for Groups 1 and 2. All three groups were treated identically on Day 2. Each subject received a total of 39, one-minute trials grouped into six blocks. Blocks 1 and 6 consisted of single-task trials only and performance on these two blocks was used as a baseline against which to measure the development of timesharing skills.

Apparatus. The stimuli for all tasks were presented on a 10.2 by 7.6 cm Hewlett-Packard Model 1300A cathode ray tube. The tracking tasks employed two identical Measurement Systems Incorporated Model 435 two-axis, spring-centered control sticks. Both sticks were modified to permit movement in the left-right dimension only. All testing was conducted in a light and sound attentuated room. Subjects were seated 116 cm from the front of the CRT for all testing. The position of the input devices (keyboards or control sticks) was adjusted for each subject.

Subjects. Sixty-five right-handed, female subjects completed two pretests, the Bennett Test of Mechanical Comprehension and one trial of the TR-TR combination. Six subjects were eliminated from the experiment because one or both pretest scores were below established criterion scores. All subjects who participated in the experiment were non-pilots and were paid an hourly wage. Monetary incentives for good performance also were given.

RESULTS

Development of Timesharing Skills

Figure 2 presents single- and dual-task tracking error as a function of practice. It is clear that with practice there is a large improvement in dual-task performance while single-task performance remains stable. Thus, the improvement in dual-task performance may be attributed to the development of timesharing skills, not to further improvements in single-task skills.

Performance on the STM-CL task combination showed a pattern similar to that of the TR-TR combination. Dual-task performance improved with practice while single-task performance remained stable. Again, these results may be interpreted as evidence for the development of timesharing skills in the discrete information processing task combination.

Transfer of Timesharing Skills

The amount of transfer of the timesharing skills from the digit task combination to the TR-TR combination was assessed by calculating the percent transfer and by examining between-group differences as a function of practice. Percent transfer was calculated based on the number of trials required to reach a criterion of 27% average absolute error of the displayed scale on both tasks simultaneously using the formula:

$$\frac{C-E}{C+E} \times 100\% = \text{percent transfer}$$

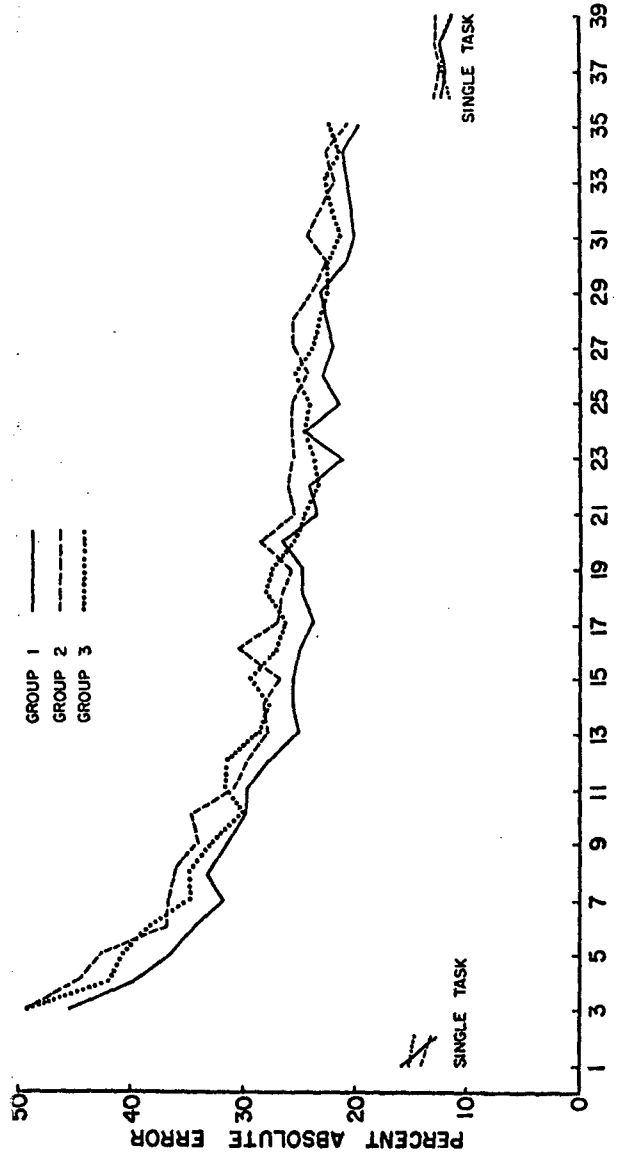


Figure 2. Average percent absolute error as a function of practice for all three groups. The unconnected data points represent single-task performance; the connected ones, dual-task performance.

where:

C is the trials to criterion for Group 3

E is trials to criterion for either Group 1 or 2

Eight subjects (three in Group 1, three in Group 2, and two in Group 3) failed to reach the 27% criterion. Group 1 had 13.9% transfer; Group 2 had ^% transfer, suggesting that the learning of the TR-TR combination benefited from dual-task but not from single-task practice on the digit task.

Between-group differences also were examined as a function of practice using analysis of covariance with the pretest scores as covariates. The analysis was conducted on the 39 subjects who reached criterion. The main effects of both trials and groups were statistically reliable ($F_{2,\ 133}$ = 115.02, $p < .001$; $F_{2,116}$ = 3.38, $p < .05$ respectively). These between-group differences indicate a reliable transfer of timesharing skills for Group 1. The transfer of these skills may be seen in Figure 1 in which it is evident that Group 1's performance is generally superior to that of Group 2 and 3. To insure that the transfer effect was not an artifact of superior single-task performance, Groups 1 and 2 were compared on their terminal Day 1 single-task performance. These scores did not differ reliably between groups.

Control Theory Analysis

The preceding results indicated that the development of timesharing skills was manifest in the reduced dual-task error with practice. It was hypothesized that two processing changes might underlie the timesharing skills responsible for this performance increase: an increase in parallel

processing or an increase in independent processing of the two information channels.

Parallel processing skill. To investigate the possibility that parallel processing increased with dual-task practice, analyses were undertaken of response holds, effective time delay, dual-task open-loop gain, and the linear coherence function. Response holds are periods of time during which the subject makes no control response although it is appropriate to do so (Cliff, 1971). If a subject develops a skill in parallel information processing, the total duration of holds in a given trial should decrease as the subject progresses from a serial to a parallel processing strategy. Furthermore, a decrease in the duration of holds should be accompanied by a corresponding reduction in the average phase lag or effective time delay between error and output (Wickens and Gopher, 1977, in press).

To investigate this possibility, the tracking control outputs of both tasks were scanned by a computer program that identified intervals of time, at least 240 msec in length, during which the output remained within a fixed amplitude window. These intervals were labelled holds and their total duration within a trial was tabulated. A four-way analysis of variance (trials, hand, group, and secondary task load) revealed no reliable effect of practice on hold duration.

The phase data then were analyzed to determine if these supported the conclusions drawn from the hold analysis. A spectral analysis (Biomedical Computer Programs, 1973) was performed on the tracking error and response data of all subjects on the first two, the middle two, and the last two dual-task trials. The phase lag data indicated, like the hold data, no change over dual-task practice and, therefore, also provide no evidence for the development of a parallel processing skill.

The emergence of parallel processing skills also was examined through the linear coherence function between error and output. A perfect parallel processing system should show a unity coherence between each error signal and the appropriate responding hand (ipsilateral coherence). As a consequence the emergence of parallel processing should be reflected by an increase in ipsilateral coherence toward the ideal unity value, even as single-task coherence remains constant. The uppper portion of Figure 3 presents the single-task coherence and the ipsilateral dual-task coherence averaged across spectral estimates (.1-2.0 Hz) and subjects at the early, middle, and late stages of practice. Single-task performance remains essentially constant while dual-task coherence for both groups increases with practice. This provides some evidence that a skill in parallel information processing developed.

This increase in linear coherence apparently is related to a corresponding increase in dual-task open-loop gain as single-task gain remains stable. This trend is shown in Figure 4 which presents a plot analogous to Figure 3 of the average amplitude ratio across spectral estimates. The characteristics of these data are almost identical to those of the coherence data.

Independent processing skill. While the extent of parallel processing may increase with practice, independent processing, the extent to which motor commands issued to each control are unaffected by simultaneous commands to the other control, also may increase. If a skill in independent information processing develops with practice, this motor "cross-talk"

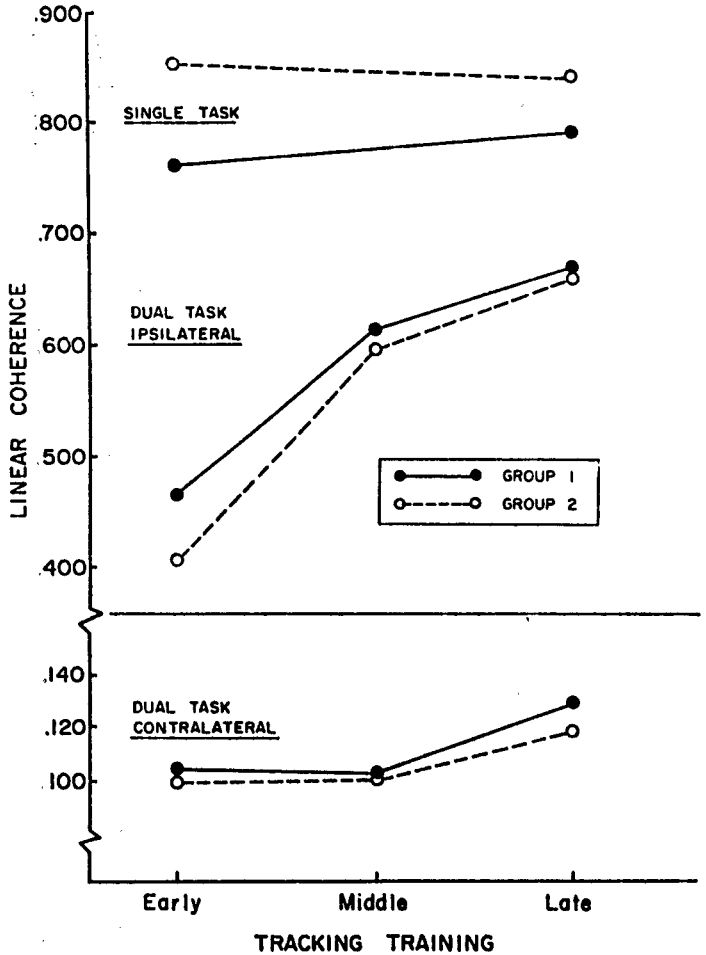


Figure 3. Linear coherence as a function of practice and secondary-task load.

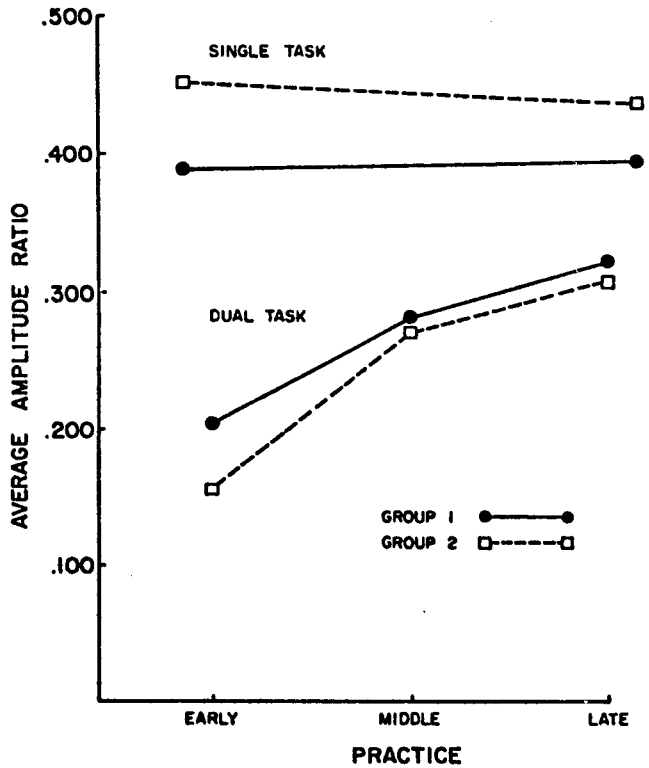


Figure 4. Average amplitude ratio as a function of practice and secondary-task load.

should decrease. To examine the extent of this crosstalk and its change with practice, the linear coherence function between each error and the opposite control response (contralateral coherence) was calculated at the three stages of practice. These data are shown in the bottom section of Figure 3. It is obvious that the contralateral coherence changes very little with practice and the change which does occur is an increase rather than a decrease.

Transfer of parallel processing skills. A fine-grained analysis of the performance of Group 1 on Day 1 revealed that some of the subjects developed a skill in parallel information processing. Because the control theory analysis indicated that a similar skill developed in the TR-TR combination, it was of interest to determine if the transfer of a parallel information processing skill between Days 1 and 2 could account for some of the transfer found using percent transfer and analysis of covariance. If such a skill transferred between the two days, the index of parallel processing shown by the subjects in Group 1 should have been initially superior to that of Group 2. With practice, however, this superiority should have diminished.

Figures 5 and 6 show the ipsilateral coherence spectrum and the amplitude ratio function as functions of practice for each group. Both measures show a change in performance with practice for both groups. More importantly, both graphs show that Group 1's performance early in practice was superior to that of Group 2, but with practice the two groups became indistinguishable. This indicates that the timesharing skill that transferred between Days 1 and 2 was, in fact, a skill in parallel information processing.

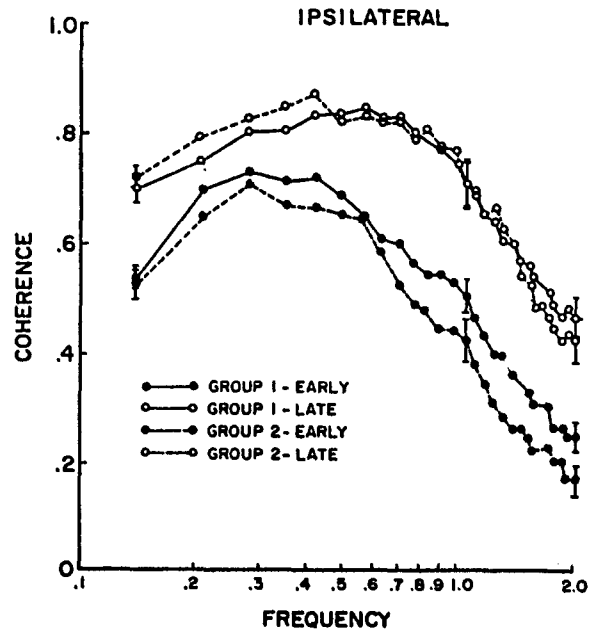


Figure 5. Ipsilateral linear coherence spectrum early and late in practice.

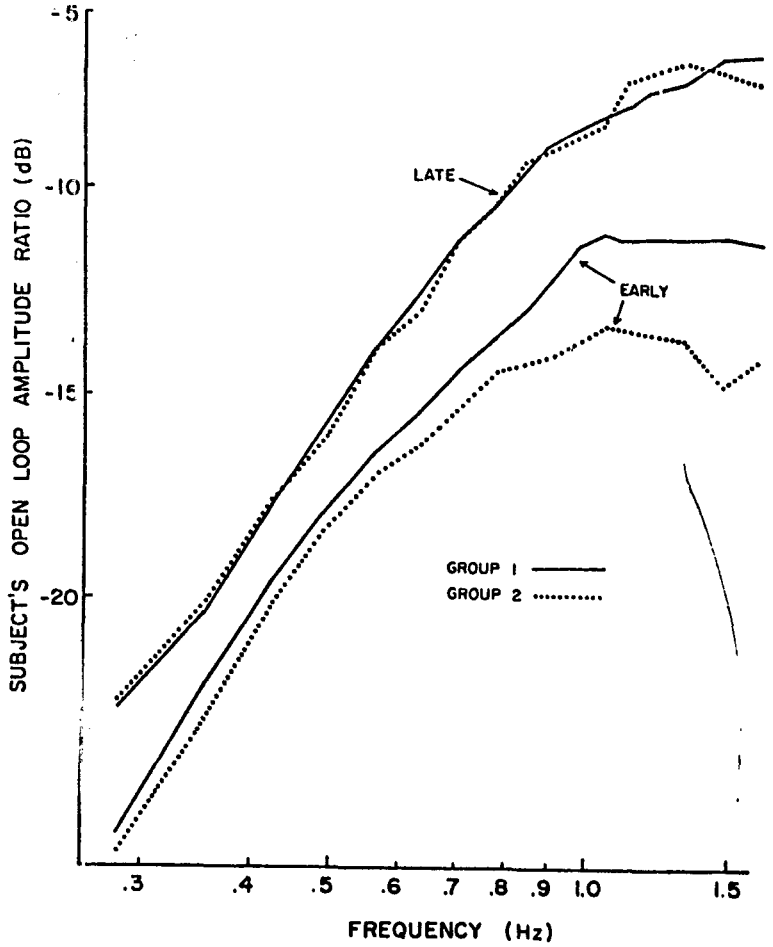


Figure 6. Amplitude ratio function early and late in practice.

SUMMARY

This experiment provides evidence for the development of timesharing skills in two qualitatively different task combinations. The calculation of percent transfer between the two combinations and an examination of between-group differences using analysis of covariance indicated that timesharing skills transfered between the combinations. A control theory analysis was performed on the TR-TR combination, the transfer task, to determine which parameters reflected the development and transfer of time-sharing skills and to identify the specific skills that developed in this combination. The ipsilateral linear coherence, open-loop gain, effective time delay, and response holds were examined for evidence of the development of a skill in parallel information processing. The development of this skill was reflected in the ipsilateral linear coherence and the open-loop gain. The contralateral linear coherence was examined for evidence of a skill in independent information processing. No evidence of such a skill was found. Examination of the ipsilateral linear coherence and the gain as a function of practice revealed that a skill in parallel information processing transfered between Days 1 and 2.

REFERENCES

Biomedical Computer Program, University of Calif. at Los Angeles, January, 1973.

Cliff, R. C. The effects of attention sharing in a dynamic dual-task environment. Proceedings of the Seventh Annual Conference on Manual Control, NASA SP-281, 1971, 307-326.

Wickens, C. and Gopher, D. Control theory measures of tracking as indices of attention allocation strategies. Human Factors, 1977, in press.

N79-17480

A Dual-Loop Model of the Human Controller

Ronald A. Hess

NASA Ames Research Center, Moffett Field, Calif.

Theme

It is the thesis of this paper[1] that a representative model of the human controller in single-axis compensatory tracking tasks should exhibit an internal feedback loop which is not evident in single-loop models now in common use[2]. This hypothetical inner-loop involves a neuromuscular command signal derived from the time rate of change of controlled element output which is due to control activity. It is not contended that the single-loop human controller models now in use are incorrect, but that they contain an implicit but important internal loop closure, which, if explicitly considered, can account for a good deal of the adaptive nature of the human controller in a systematic manner.

Contents

Figure 1 is a block diagram representation of the hypothetical human controller model. This is referred to as the dual-loop model to distinguish it from typical single-loop structures. The dual-loop model is not derived from any physiological imperatives but rather stems from a structurally similar but philosophically different model discussed briefly by Smith[3]. Table 1 presents the parameters associated with the model of Fig. 1. The equalization Y_{p_e} contains a simple first-order lead term and a high frequency lag term with break frequency $1/T_e$ beyond the undamped natural frequency ω_n of the neuromuscular system. The exponent 'a' on the lag is included to implement more rapid fall-off in the filtering characteristics of Y_{p_e}, if necessary.

The equalization Y_{p_m} consists of a low frequency washout which, in terms of an acceptable feedback control system, is essential for inner-loop low frequency remnant suppression. Inner-loop low frequency remnant power can arise from a variety of sources including errors in the controller's internal model of manipulator/controlled element dynamics $\hat{Y}_\delta Y_c$ at low frequencies, poor kinesthetic feedback u_δ at low frequencies or broadband process noise. Just as in Y_{p_e}, the integer exponent 'b' can be used to implement sharper washout characteristics, if necessary. As Table 1 indicates, the $\hat{Y}_\delta Y_c$ element represents the human controller's internal model of the manipulator/controlled element dynamics. Indeed, the dominant adaptive feature of the dual-loop model is the explicit appearance of $\hat{Y}_\delta Y_c$ in the equivalent single-loop form of the model in Table 1 (Y_p).

The element Y_{p_n} represents the neuromuscular dynamics of the particular limb which drives the manipulator. The dynamics shown in Table 1 have been deliberately simplified from the more elaborate neuromuscular model presented in ref. 2 for the sake of simplicity. The remnant signals r_e and $r_{\dot{u}_\delta}$ are injected into displayed error e_d and rate of change of control force $\dot{u}_\delta$. These remnant signals are part of the quasi-linear representation of the outer and inner-loop human controller characteristics.

Implicit in the dual-loop formulation is the assumption that the structure as outlined in Table 1 is complete, i.e., no additions have to be made or restructuring undertaken to account for human controller adaptation to various controlled elements, displays, manipulators, etc. The adaptive potential of the model is contained in the parameters of the inner and outer-loop equalization and in the internal manipulator/controlled element model $\hat{Y}_\delta Y_c$. Four hypotheses

regarding the general adaptive characteristics of the parameterized dual-loop model are offered: Hypothesis 1 - The parameters K_e, $K_{\dot{m}}$, and $T_{\dot{m}}$ act in consort with the internal model of the manipulator/controlled element dynamics $\hat{Y}_\delta Y_c$ to define the essential adaptive capabilites of the dual-loop model. The lead equalization in Y_{p_e} is used only when the previous parameters are unable to provide the effective lead equalization which may be required in a specific task. Hypothesis 2 - The relative utilization of the inner and outer-loops of the dual-loop model is a function of manipulator/controlled element dynamics and the quality of the sensory inputs e_d and u_δ. Hypothesis 3 - Relative loop utilization in the dual-loop model can be quantified by the ratio $K_{\dot{m}}/K_e$. Hypothesis 4 - The value of $T_{\dot{m}}$ is determined by the extent to which the inner-loop is utilized and by the quality of the internal model of the manipulator/controlled element dynamics $\hat{Y}_\delta Y_c$. The more the inner-loop is utilized, the larger the value of $T_{\dot{m}}$. The more precise the internal model, the larger the value of $T_{\dot{m}}$. No hypotheses regarding the adaptive nature of the parameters T_e (or τ), ζ_n and ω_n are offered. These parameters effect primarily the high frequency portion of the controller describing function Y_p ($\omega > 10$ rad/sec). Since measured data tends to be scant in this frequency range, verifying hypotheses regarding parameter variations would be difficult.

The dual-loop model just outlined can produce describing functions which closely approximate those measured in laboratory tracking tasks. The tasks used in the initial validation involved the following controlled element dynamics: K, K/s, $K/(s)^2$, $K/(s^3+12.3s^2+11.6s)$ and $K/s(s-1)$. In addition, an empirical remnant model utilizing a pair of dual-loop model parameters can approximate the measured error remnant spectra for those experiments for which error injected remnant data were reported (all but the fourth task above).

No formal numerical algorithm was utilized in identifying the dual-loop model parameters for the five experimental tasks. Instead, small and large bandwidth approximations to the last equation in Table 1 were derived and a straight-forward hand-fit procedure was employed for parameter selection and adjustment. The pertinent expressions for Y_p are, for small bandwidth

$$Y_p = u_\delta/e_d \doteq \frac{K_e(T_{\dot{m}}s+1)(T_L s+1)}{(K_{\dot{m}}\hat{Y}_\delta Y_c)s^2+(T_{\dot{m}})s+1} \qquad (1)$$

and for large bandwidth

$$Y_p = u_\delta/e_d \doteq \frac{K_e(T_{\dot{m}}s+1)(T_L s+1)e^{-\tau s}}{(T_{\dot{m}}/\omega_n^2)s^3+(2\zeta_n T_{\dot{m}}/\omega_n+K_{\dot{m}}\hat{Y}_\delta Y_c+1/\omega_n^2)s^2+(T_{\dot{m}}+2\zeta_n/\omega_n)s+1} \qquad (2)$$

Relations (1) and (2) and the general guidelines of the four hypotheses just stated, allow rapid "identification" of the dual-loop model parameters using measured describing function data through the following two-step process. First, K_e, $K_{\dot{m}}$, $T_{\dot{m}}$ and, if necessary, T_L are selected using equation (1) and the low to mid-frequency ($\omega < \omega_c$ where ω_c is the open-loop crossover frequency) describing function data. Second, ζ_n, ω_n and τ are chosen using equation (2), the selected values of K_e, $K_{\dot{m}}$, $T_{\dot{m}}$ and T_L, and the high frequency ($\omega > \omega_c$) describing function data.

An empirical expression for the normalized power spectral density of the "equivalent" remnant which is injected into the displayed error of the dual-loop model can be obtained from the measured remnant data as

$$\Phi_{nn_e}/\sigma_e^2 = \frac{P_e(K_{\dot{m}}/K_e)}{1+(K_{\dot{m}}/K_e)^2\omega^2/2} \qquad (3)$$

where σ_e is the root-mean-square value of the displayed error signal and

P_e is a parameter to account for the fact that different levels of human controller attention are required for different control tasks. Only two values of P_e are needed to match the measured remnant data in this study: $P_e = 0.2$ to represent a "low" attention level and $P_e = 0.04$ to represent a "high" attention level.

The dual-loop model can explain the effects of specific display and manipulator variations on human controller characteristics. For example, Fig. 2 shows a comparison between measured and model generated describing functions and remnant spectra for a tracking task with K/s controlled element dynamics and a peripheral display. The match is quite acceptable. In particular, the remnant model does an excellent job of matching the experimental remnant spectrum in spite of the fact that the former was empirically derived from foveal viewing data.

The assessment of a dual-loop model of the human controller undertaken in this study leads to the following conclusions: (1) The proposed dual-loop model and the general adaptive characteristics which have been hypothesized can produce human controller describing functions which closely approximate those measured in a wide variety of single-axis compensatory tracking tasks. (2) An empirical model for injected remnant spectra employing low and high controller attention levels can approximate experimentally derived injected remnant spectra. Of the experiments to which the remnant models were applied, the low attention model satisfactorily matched the data for the stable controlled elements while the high attention model matched the data for the unstable elements. (3) The dual-loop model and associated hypotheses can explain the measured variations in human controller dynamics and performance which accompany changes in controlled element dynamics and variations in display and manipulator characteristics. (4) In terms of existing single-loop models, the dual-loop model exhibits the following novel features: (a) the adaptive nature of the model is due primarily to the existence of an explicit internal model of the manipulator/controlled element dynamics in an inner feedback loop, (b) in controlling all but the most difficult controlled element, the dual-loop model indicates that the necessity of lead equalization in the form of error rate utilization is obviated, i.e., in all but one case, $T_L = 0$. Likewise, apparent error lag equalization for pure gain controlled elements is accomplished by inner-loop activity.

References

[1]Hess, R. A., "A Dual-Loop Model of the Human Controller in Single-Axis Compensatory Tracking Tasks," NASA TM-73,249, May 1977

[2]McRuer, D. T., "Development of Pilot-in-the-loop Analysis," Journal of Aircraft, Vol. 10, Sept. 1973, pp. 515-524.

[3]Smith, R. H., "A Unified Theory for Pilot Opinion Rating," Proceedings of the Twelfth Annual Conference on Manual Control, NASA TM X-73,170, May 1976, pp. 542-558.

Table 1. Dual-Loop Model Parameters

Model Element	Form	Description
Y_{p_e}	$\frac{K_e(T_L s+1)}{(T_e s+1)^a}$	displayed error equalization a = 1,2, . .
$Y_{p_{\dot{m}}}$	$\frac{K_{\dot{m}} s^b}{(T_{\dot{m}} s+1)^b}$	control rate equalization b = 1,2, . .
$\hat{Y}_\delta \hat{Y}_c$		human controller's internal model of manipulator/controlled element dynamics
Y_{p_n}	$\frac{1}{(s/\omega_n)^2+(2\zeta_n/\omega_n)s+1}$	neuromuscular dynamics
r_e, $r_{\dot{u}_\delta}$		remnant injected into displayed error and control rate, respectively
$Y_p = u_\delta/e_d$	$\frac{Y_{p_e} Y_{p_n}}{1+Y_{p_n} \hat{Y}_\delta \hat{Y}_c Y_{p_{\dot{m}}} s}$	equivalent single-loop human controller describing function

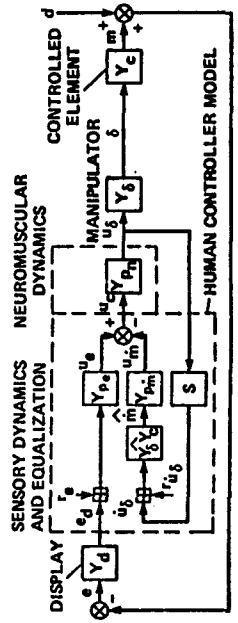


Figure 1. Quasi-Linear Dual-Loop Model of the Human Controller in a Single-Axis Compensatory Tracking Task

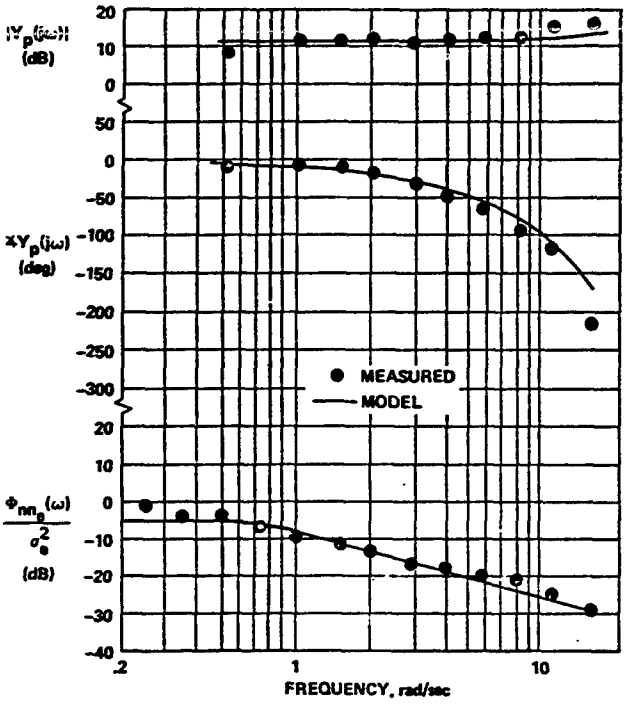


Figure 2. Comparison Between Measured and Dual-Loop Describing Functions and Injected Remnant Spectra for K/s Dynamics and 22° Peripheral Viewing

N79-17481

PARAMETER ESTIMATION IN A HUMAN OPERATOR DESCRIBING FUNCTION MODEL FOR A TWO-DIMENSIONAL TRACKING TASK.

Antonie van Lunteren

Man-Machine Systems Group
Laboratory for Measurement and Control
Department of Mechanical Engineering
Delft University of Technology
Mekelweg 2
Delft
Netherlands

o Summary

A previously described parameter estimation program has been applied to a number of control tasks, each involving a human operator model consisting of more than one describing function.
One of these experiments is treated in more detail. It consisted of a two-dimensional tracking task with identical controlled elements.
The dynamics of these controlled elements have been chosen as K, K/s, and K/s^2. The tracking errors were presented on one display as two vertically moving horizontal lines. Each loop had its own manipulator. The two forcing functions were mutually independent and consisted each of 9 sine waves.
A human operator model was chosen consisting of 4 describing functions, thus taking into account possible linear crosscouplings. From the Fourier coefficients of the relevant signals the model parameters were estimated after alignment, averaging over a number of runs and decoupling.
The results show that for the elements in the main loops the crossover model applies. A weak linear cross-coupling existed with the same dynamics as the elements in the main loops but with a negative sign.

1 Introduction

In an earlier paper [1] a parameter estimation method has been treated which gives consistent estimates in closed loop systems. This is an essential requirement for identification of human operator models in tracking experiments. The method has been based on the application of forcing functions consisting of a number of sinusoids. It has been shown that, after application of a decoupling procedure, it can also be applied for human operator models with more than one input and more than one output. The paper just mentioned does not give applications of the method. Since then, however, the method has been applied in a number of experiments.

Because much knowledge is already available on human operator behavior in single loop compensatory tracking tasks, applications have been chosen involving more complex human operator describing function models. These applications were:

- a multiple loop control task,
- a multiloop control task,
- a control task with preview.

They have been described in a recent Progress Report of the Man-Machine Systems Group [2]. This paper deals with one of the three topics just-mentioned viz the multiple loop control task.

2 Multiple loop control tasks

A multiple loop control task consists of controlling a number of mutually independent systems. Fig. 1 gives the block diagram for the case of a multiple loop task with two control loops. The most general representation of a human operator describing function model for this type of task is given in Fig. 2. In the ideal case the describing functions 21 and 12 would be zero. However, as a result of task interference these describing functions may have a nonzero value.

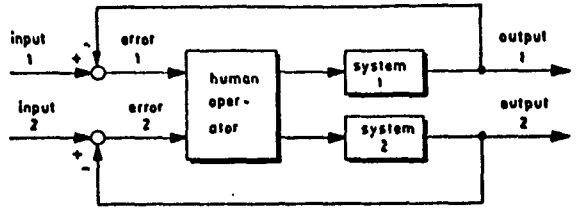


Figure 1: A multiple loop control task with two control loops.

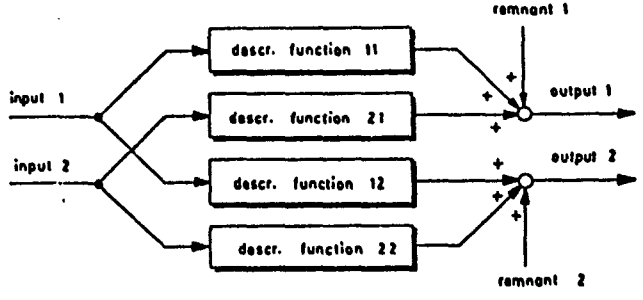


Figure 2: Human operator model with two inputs and two outputs.

Since 1960 a number of attempts to investigate multiple loop systems have been made by Chernikoff et al [3,4,5] using two predictable forcing functions, each consisting of a single sunusoid. As a measure for the human operator performance only the error scores were determined. Investigations resulting in human operator describing functions in a multiple loop control task date from about 1965 and have been published by Todosiev et al [6] and by Levison [7,8]. Several experimental parameters were varied, like system dynamics (K,K/s,K/s^2, equal or different systems) control-display configurations (one two-dimensional display and control; two one dimensional displays and controls, both signals foveally visible or not) and forcing function bandwidth. In all studies no a priori assumption was made that linear crosscouplings due to task interference might exist within the human operator model. The absence of such an assumption greatly facilitates the data processing procedure.
In the present study the possible occurence of linear cross-couplings due to task interference was taken into account, so a model structure according to Fig. 2 was assumed. To eliminate effects due to scanning behavior as well as possible it was decided to display both tracking errors foveally on one CRT. A display-control configuration with a good compatibility was desired in order to prevent effects resulting from suboptimal task conditions. Two possibilities exist: Application of one joystick which can be moved in left-right direction for one control loop and in a forward-backward direction for the other. The tracking errors can be indicated by one lightspot which can move horizontally and vertically (integrated configuration). The second possibility is to use two joysticks which can be moved in a forward-backward direction with the left hand and the right hand respectively. The corresponding tracking errors can be displayed as two vertically moving dots on the left hand and right hand part of the CRT (separated configuration). In the integrated configuration differences in performance may be found in the horizontal and in the vertical direction. In the separated configuration left-right differences may occur resulting from left- or righthandedness of the subject. According to Levison and Elkind [8] task interference may easily occur in an integrated configuration with a heterogeneous control task (each loop a different controlled element). Although the present study was limited to homogeneous control tasks (both loops the same controlled element) the separated control configuration was chosen.

3 *System identification*

In the introduction it has already been mentioned that the method applied has been treated in an earlier paper [1]. Therefore only the main points will be summarized here. For more details see also [9]. When applying estimation procedures it is always important to have a measure for the reliability of the results. Therefore a method has been developed to compute bias and standard deviation for estimates in closed loop systems based on a finite observation time. An extensive treatment is given in an internal report [10] (in Dutch). Here only the resulting expressions are given.

3.1 *Identification method*

Consider an unknown system with transfer function $H_1(\nu)$ (ν=frequency in Hz) in a closed loop disturbed by noise. The configuration is given in Fig. 3. Suppose that the external forcing function u(t) can be chosen as a

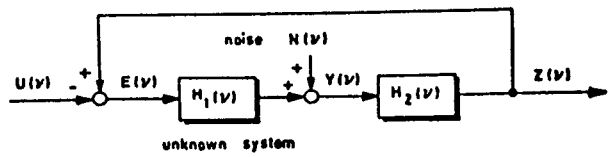


Figure 3: Unknown system in a closed loop disturbed by noise.

sum of sinusoids and that the error signal e(t) and the disturbed output y(t) of the unknown system can be measured. If the transfer function $H_1(\nu)$ can be characterized by the relation

$$H_1(\nu) = \frac{B(\nu)}{A(\nu)} e^{-j2\pi\nu\tau} , \tag{1}$$

with

$$B(\nu) = b_0+b_1(j2\pi\nu)+\ldots+b_M(j2\pi\nu)^M , \tag{2}$$

$$A(\nu) = a_0+a_1(j2\pi\nu)+\ldots+a_L(j2\pi\nu)^L , \tag{3}$$

then the following relation exists:

$$A(\nu)Y(\nu)-B(\nu)e^{-j2\pi\nu\tau}E(\nu) = A(\nu)N(\nu) . \tag{4}$$

The term $A(\nu)N(\nu)$ is called the system residue $V(\nu)$. The signals $Y(\nu)$ and $E(\nu)$ each can be split up in two parts resulting from the forcing function $U(\nu)$ and the noise $N(\nu)$ respectively, thus:

$$E(\nu) = E_u(\nu) + E_n(\nu) , \tag{5}$$

$$Y(\nu) = Y_u(\nu) + Y_n(\nu) . \tag{6}$$

For the undisturbed signals $E_u(\nu)$ and $Y_u(\nu)$ the following relation holds:

$$A(\nu_i)Y_u(\nu_i)-B(\nu_i)e^{-j2\pi\nu_i\tau}E_u(\nu_i) = 0 \tag{7}$$

where ν_i are the frequencies for which the forcing function $U(\nu)$ contains a sinusoidal component (i=1...n).
From the measured signals e(t) and y(t) estimates $\hat{E}_u(\nu_i)$ and $\hat{Y}_u(\nu_i)$ can be made for $E_u(\nu_i)$ and $Y_u(\nu_i)$ by means of fourier analysis. If the system is modelled by a transfer function

$$\hat{H}_1(\nu) = \frac{\hat{B}(\nu)}{\hat{A}(\nu)} e^{-j2\pi\nu\hat{\tau}} , \tag{8}$$

then for this model the following equation holds:

$$\hat{A}(\nu_i)\hat{Y}_u(\nu_i)-\hat{B}(\nu_i)e^{-j2\pi\nu_i\hat{\tau}}\hat{E}_u(\nu_i) = \hat{V}(\nu_i) , \; i=1,..n . \tag{9}$$

In this equation the quantity $\hat{V}(\nu_i)$ is called the model residue at frequency ν_i. The model parameters are estimated by minimizing the following cost function:

$$J = \sum_{i=1}^{n} \lambda_i |\hat{V}(\nu_i)|^2 \ . \qquad (10)$$

The quantities λ_i are weighting factors. For a given value τ of the delay time $\hat{\tau}$ to be estimated, minimization of the cost function J to the unknown parameters $\hat{a}_j$(j=0,1,..L) and $\hat{b}_k$(k=0,1,..M) leads to a set of linear equations in these unknown parameters. Without loss of generality one of these parameters, for instance a_o can be chosen as equal to 1. Substitution of these solutions in the original equation and differentiation with respect to the model parameter $\hat{\tau}$ leads to one non-linear equation from which the parameter τ can be solved, after which also the other parameters can be calculated.

For the estimator of the parameters $\hat{a}_j$ and $\hat{b}_k$ the following properties can be derived [9]:
- The estimator is consistent.
- The estimator will be a minimum variance estimator if the weighting factors are chosen as: $\lambda_i = S_{vv}(\nu_i)$, where $S_{vv}(\nu_i)$ is the auto spectral density of the system residue v(t) at the frequency ν_i. According to (4) this quantity can also be given by

$$\lambda_i = S_{vv}(\nu_i) = |A(\nu_i)|^2 S_{nn}(\nu_i) \ . \qquad (11)$$

In practice both $A(\nu_i)$ and $S_{nn}(\nu_i)$ are unknown. A possible procedure is to make a first estimate with weighting factors λ_i=1, based on these estimates, to obtain an estimate for the weighting factors. These factors are used in a new estimate. If necessary this procedure can be repeated.

3.2 *Reliability of the estimates*

Consistency of an estimator means that for an infinite observation time the estimate will become equal to the quantity to be estimated with probability 1. For a finite observation time the estimator will have a certain variance and may even have a bias, i.e. a consistent estimator may only be asymptotically unbiased. It has been shown elsewhere [10] that in the case of Fig. 3 a system in the loop can only be identified by an *asymptotically* unbiased estimator. For a finite observation time, for a normally distributed remnant, and for a forcing function u(t) described by:

$$u(t) = \sum_{i=1}^{n} a_i \cos 2\pi\nu_i t + b_i \sin 2\pi\nu_i t \qquad (12)$$

the bias in an estimate $\hat{H}_1(\nu_i)$ of $H_1(\nu_i)$ is equal to:

$$\text{Bias}\{H_1(\nu_i)\} = E\{\hat{H}_1(\nu_i) - H_1(\nu_i)\} = -\{H_1(\nu_i) + \frac{1}{H_2(\nu_i)}\} e^{-x_i} \ , \qquad (13)$$

where

$$x_i = \left|\frac{E_u(\nu_i)}{E_n(\nu_i)}\right|^2 = \left|\frac{\frac{2}{T}\int_0^{kT} e_u(t) e^{-j2\pi\nu_i t} dt}{\frac{2}{T}\int_0^{kT} e_n(t) e^{-j2\pi\nu_i t} dt}\right|^2 \qquad (14)$$

and where k is an integer and T is the basic period of the periodic test signal u(t). Thus the quantity x_i can be considered as the signal to noise ratio of the signal e(t) at the frequency ν_i. This ratio increases with increasing value of the integer k belonging to the observation time kT.
For the variance of the estimator $\hat{H}_1(\nu_i)$ the following expression has been derived:

$$\text{Var}\{\hat{H}_1(\nu)\} = E\{|\hat{H}_1(\nu_i) - E\{\hat{H}_1(\nu_i)\}|^2\} = |H_1(\nu_i) + \frac{1}{H_2(\nu_i)}|^2 (1 + e^{-2x_i} - x_i e^{x_i} E_1(x_i)), \qquad (15)$$

where the function $E_1(x)$ is defined as:

$$E_1(x) = \int_x^{\infty} \frac{e^{-v}}{v} dv \ . \qquad (16)$$

The functions $E_1(x)$ and $xe^x E_1(x)$ are given in tables, for instance in Abramowitz and Stegun [11].
If the input u(t) is stochastic the expressions for bias and variance are almost identical. Only the quantity x_i must be replaced by $x_n(\nu)$, where

$$x_n(\nu) = \frac{n\{\Gamma_{ue}(\nu)\}^2}{1 - \{\Gamma_{ue}(\nu)\}^2} \ . \qquad (17)$$

In this expression $\Gamma_{ue}(\nu)$ is the coherence between the signals u(t) and e(t) at the frequency ν; n is the number of elementary frequency bands with width $\Delta\nu = 1/T$, over which an averaging is executed in the estimation procedure for the spectral densities $S_{ue}(\nu)$ and $S_{ue}(\nu)$ [12]. These estimates are used to obtain the transfer function estimate:

$$\hat{H}_1(\nu) = \frac{\hat{S}_{uy}(\nu)}{\hat{S}_{ue}(\nu)} \ . \qquad (18)$$

A study of the expressions for bias and variance shows that the bias decreases much more rapidly with increasing observation time, so that in many practical cases the bias can be neglected.

3.3 *Complex configurations*

For systems with more than one input and more than one output a number of separate transfer functions has to be estimated. In such a case (see for instance Fig. 2) more than one external forcing function has to be applied and a decoupling procedure has to be executed on the available inputs and outputs. This procedure has already been described earlier [1,13] It involves a number of subtractions and divisions of fouriercoefficients of the measured inputs and outputs. As a consequence thereof it is no longer possible to get reliable estimates for bias and variance as in the simple case of Fig. 3.

4 *Experiments and data processing.*

The control configuration is given in the block diagram of Fig. 4. The

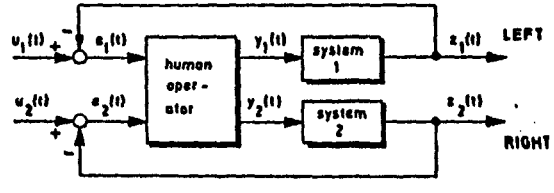


Figure 4: The control configuration for the multiple loop task.

forcing functions $u_1(t)$ and $u_2(t)$ each consisted of 9 sinusoids each having an integer number of periods in a 102.4 seconds time interval. For each of the components the number of periods in this interval is given in Table 1.

Table 1: Number of periods in 102.4 sec for the components of the forcing functions $u_1(t)$ and $u_2(t)$.

comp. nr.	1	2	3	4	5	6	7	8	9
u_1	3	5	9	16	28	51	97	179	308
u_2	4	6	10	17	29	52	98	180	309

The signals $u_1(t)$ and $u_2(t)$ were derived from signals with a constant amplitude spectrum, which were sent through a first order low-pass filter with a time constant of 0.43 sec (break point at 0.37 Hz). The error signals $e_1(t)$ and $e_2(t)$ were presented on a 10x8 cm CRT as two vertically moving dots, horizontally 1 cm apart, at a distance of 50 cm from the subject's eyes. The sensitivity of the CRT was 1V/cm. The subject had two identical fingertip control sticks, one for each hand, which could be moved in a forward-backward direction; the stick gain was 40 V/cm. The stiffness was 0.04 N/cm. In all experiments the systems 1 and 2 were identical. For these systems three different elements were chosen:

$$H_{c1}(\nu) = H_{c2}(\nu) = 1;$$

$$H_{c1}(\nu) = H_{c2}(\nu) = 4/j2\pi\nu ;$$

$$H_{c1}(\nu) = H_{c2}(\nu) = 8/(j2\pi\nu)^2 .$$

The experiments were carried out with three right handed male students. They were trained with each of the systems, single loop right hand, single loop left hand and both loops simultaneously until their mean square error reached a stationary value.
Then the final experiments were executed according to the program of Table 2. The duration of one run was about 3 minutes, of which an interval of 102.4 sec was used for data analysis. Therefore the eight signals indicated in Fig. 4 were recorded on magnetic tape together with a 160 Hz clock signal to ensure analysis over an integer number of periods for each of the components of the forcing function.

Table 2 : Schedule of the experiments.

date	system	hand	runs per subject
1973 05-22	1	L + R	4
		R	1
		L	1
1973 05-23	$4/j2\pi\nu$	L + R	4
		R	1
		L	1
1973 05-24	$8/(j2\pi\nu)^2$	L + R	4
		R	1
		L	1

The signal analysis consisted for each signal of:
- determination of the standard deviation,
- transformation by means of the FFT,
- selection of the Fourier coefficients at the 18 forcing function frequencies,
- determination of the ratio between the variances of the remaining noise part of the signal and of the complete signal.

In order to obtain a more reliable estimate of the average control behavior of the three subjects together, an averaging procedure over the estimated Fourier coefficients was applied. This procedure includes a rotation of the vectors represented by the sine and cosine coefficients in the complex plane. In fact this rotation aligns the signal components with respect to a common zero reference in time. The averaged Fourier coefficients thus obtained were used as input for the parameter estimation program.

5. *Results.*

Table 3 shows the averaged signal data over the experiments. Averaging was done over the 3 subjects and over all available runs for each subject (see Table 2). For the 12 different experimental conditions (3 systems, left-right, single or double task) the means μ and standard deviations σ are given for the estimates of the standard deviations σ_u, σ_e, σ_y, σ_z respectivily of the input signal $u(t)$, the error signal $e(t)$, the human operator output $y(t)$ and the system output $z(t)$. For these four signals the means μ and the standard deviations σ of the relative noise contents N_u, N_e, N_y and N_z were estimated. The relative noise content is defined as the ratio between on the one hand the variance of that part of the signal which is located at frequencies for which no components are present in the forcing function $u(t)$ and on the other the variance of the complete signal. Finally , the normalized tracking error σ_e/σ_u is given as a measure for the performance of the subjects. Due to limitations in the capacity of the tape recorder the signals $u_2(t)$ and $z_2(t)$ were not processed for the two handed task, therefore quantities from these signals have been left open in the table.

Table 3 : Mean value μ and standard deviation σ of the estimates of the standard deviations (in cm display) and relative noise content of the four signals u(t), e(t), y(t) and z(t), and the relative tracking error σ_e/σ_u.

system	1				$4/(j2\pi\nu)$				$8/(j2\pi\nu)^2$			
hand	left		right		left		right		left		right	
task	single	double	single	double	single	double	single	double	single	double	single	double
σ_u	1.49	1.23	1.36		1.03	1.05	1.04		0.3	0.3	0.3	
	0.00	0.00	0.00		0.08	0.07	0.08		0.1	0.1	0.1	
σ_e	0.56	0.56	0.55	0.57	0.55	0.62	0.51	0.63	0.5	0.6	0.5	0.6
	0.02	0.04	0.01	0.05	0.01	0.02	0.02	0.03	0.1	0.1	0.1	0.1
σ_y	1.11	1.10	1.09	1.10	1.1	0.7	0.8	0.7	1.2	0.8	1.1	1.1
	0.05	0.06	0.08	0.06	0.3	0.1	0.2	0.1	0.5	0.5	0.5	0.5
σ_z	1.14	1.14	1.12		1.11	1.13	1.12		0.13	0.17	0.15	
	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	
N_u	0.000	0.000	0.000		0.000	0.000	0.000		0.000	0.000	0.000	
	0.000	0.002	0.001		0.002	0.002	0.002		0.000	0.000	0.000	
N_e	0.15	0.15	0.17	0.15	0.24	0.27	0.21	0.22	0.08	0.01	0.01	0.03
	0.06	0.05	0.01	0.05	0.09	0.06	0.10	0.07	0.02	0.02	0.04	0.05
N_y	0.035	0.041	0.032	0.035	0.4	0.2	0.5	0.3	0.1	0.1	0.1	0.1
	0.007	0.014	0.003	0.022	0.3	0.15	0.1	0.1	0.1	0.1	0.1	0.1
N_z	0.031	0.041	0.035		0.16	0.06	0.09		0.07	0.05	0.10	
	0.007	0.008	0.008		0.09	0.03	0.03		0.02	0.03	0.09	
$\frac{\sigma_e}{\sigma_u}$	0.38	0.45	0.40		0.53	0.59	0.49		1.6	2.3	1.7	
	0.01	0.00	0.00		0.10	0.10	0.09		0.1	0.1	0.1	

The table shows that the standard deviation σ_e of the error signal e(t) was always about 0.6 cm. In fact, using information from the training period, the inputs $u_1(t)$ and $u_2(t)$ were chosen in such a way that σ_e was more or less the same during all experiments. The relative noise content N_u of the input signals $u_1(t)$ and $u_2(t)$ was always very low (<0.01). This is an indication that the signal/noise ratio of the recorder was sufficiently low, and that the fourier analysis was indeed based on an observation time equal to an integer number of periods for all input frequencies. As might be expected the relative tracking error increases with increasing system order. It looks as if in a single task the left hand performed slightly better than the right hand although all three subjects were righthanded. Of the same order are the differences which give an indication for a performance deterioration in the double task compared with the single task.

Fig. 5 shows examples of the amplitude spectra for the three different systems in a double task. The figures on the relative noise content in Table 3 already indicate an increase of human operator remnant with increasing system order. The spectra, moreover, inllustrate how the remnant is distributed along the frequency axis.

For the experiments with two control loops the mathematical description of the man-machine system is illustrated in the block diagram of Fig. 6 . The human operator model transfer functions were identified as follows: From the averaged and decoupled fouriercoefficients the raw data points for the Bode plots of the four transfer functions $H_{11}(\nu)$, $H_{12}(\nu)$, $H_{21}(\nu)$ and $H_{22}(\nu)$ were determined. These data points are given by the crosses in Fig. 7 . Based on these data points models were chosen for the four transfer functions for each of the three experimental conditions. By means of the parameter estimation program the parameters in these models were estimated from the fouriercoefficients of the decoupled signals. Using these parameter values the model transfer functions were obtained and also drawn in Fig. 7 (solid lines). The same procedure was applied to the single loop measurements, be it without the decoupling operation. The resulting models with their parameter values are given in Table 4.
It is interesting to note that linear cross couplings were clearly present in the cases that two proportional systems or two integrators had to be controlled. All cross couplings had a negative sign. In the case of the two double integrators there was a cross coupling from the right hand display (e_2) to the left hand control (y_1), again with a negative sign. From the left hand display to the right hand control no significant linear coupling could be found. Attempts to fit a model in the latter case yielded modelling errors in the order of 90%, while in all other cases these values were in the order of 10% or lower.
The raw data points and the Bode plots of the human operator models for the single loop experiments are given in Fig. 8.
In the averaging procedure for the fouriercoefficients not only the measurements but also the standard deviations were determined. In this way estimates for the signal to noise ratio for a number of frequencies were available. Using these values it was possible to estimate bias and standard deviation according to Eqs. 13 and 15 from section 3.2. The resulting one sigma limits are also indicated in Fig. 8.

From the results given in Table 4 the cross-over frequencies and phase margins were determined for the single loop tasks and for the comparable main loops in the double loop tasks. The results are given in Table 5.

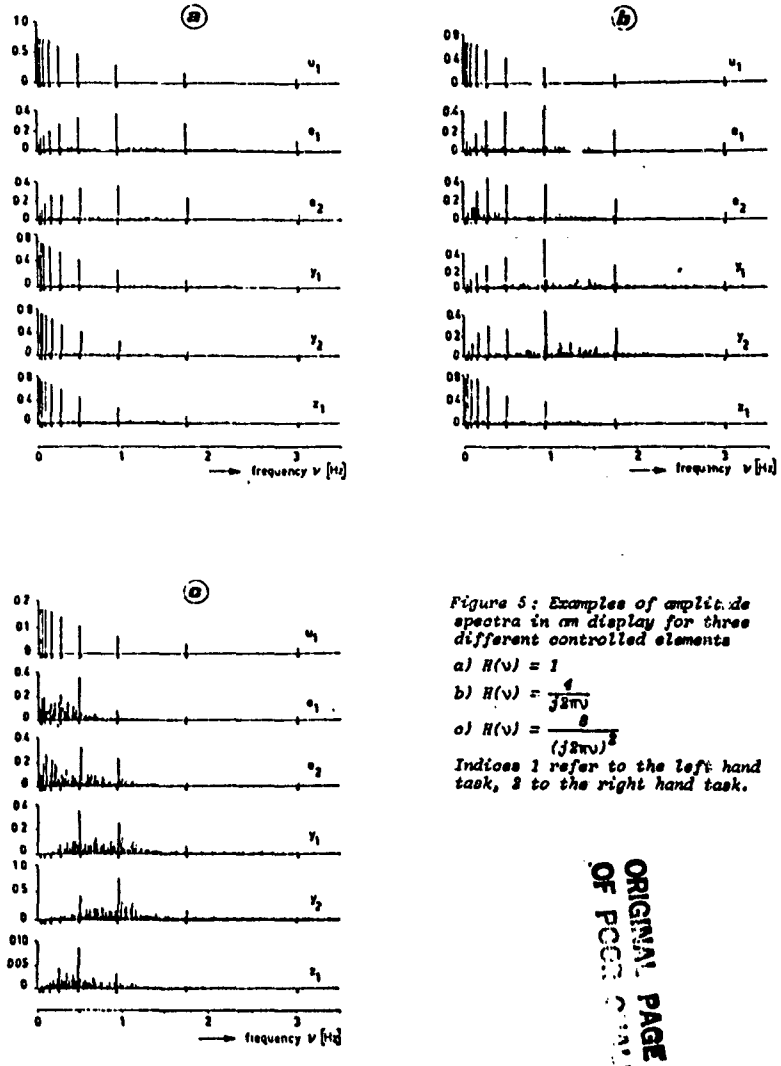


Figure 5: Examples of amplitude spectra in cm display for three different controlled elements

a) $H(\nu) = 1$

b) $H(\nu) = \frac{1}{j2\pi\nu}$

c) $H(\nu) = \frac{1}{(j2\pi\nu)^2}$

Indices 1 refer to the left hand task, 2 to the right hand task.

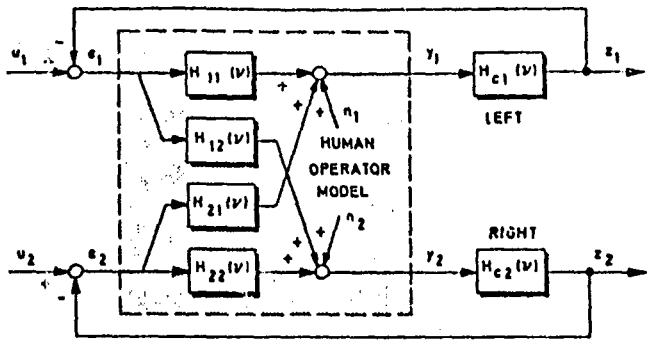


Figure 6: Block diagram of the human operator model in a two dimensional multiple loop system.

Table 4: Human operator describing functions.

$H_c(\nu)$	hand	single system	
1	LL	$\frac{10}{1+3(j2\pi\nu)}\,e^{-0.11(j2\pi\nu)}$	$\frac{8}{1+2.5(j2\pi\nu)}\,e^{-0.12(j2\pi\nu)}$
	LR		$-\frac{0.2}{j2\pi\nu}\,e^{-0.22(j2\pi\nu)}$
	RL		$-\frac{0.5}{j2\pi\nu}\,e^{-0.13(j2\pi\nu)}$
	RR	$\frac{6}{1+1.8(j2\pi\nu)}\,e^{-0.10(j2\pi\nu)}$	$\frac{6}{1+2(j2\pi\nu)}\,e^{-0.12(j2\pi\nu)}$
$\frac{1}{j2\pi\nu}$	LL	$\frac{0.2+2.2(j2\pi\nu)}{j2\pi\nu}\,e^{-0.15(j2\pi\nu)}$	$\frac{0.2+1.2(j2\pi\nu)}{j2\pi\nu}\,e^{-0.18(j2\pi\nu)}$
	LR		$-0.08\,e^{-0.26(j2\pi\nu)}$
	RL		$-0.10\,e^{-0.11(j2\pi\nu)}$
	RR	$\frac{0.3+1.1(j2\pi\nu)}{j2\pi\nu}\,e^{-0.17(j2\pi\nu)}$	$\frac{0.2+0.9(j2\pi\nu)}{j2\pi\nu}\,e^{-0.20(j2\pi\nu)}$
$\frac{1}{(j2\pi\nu)^2}$	LL	$0.4(j2\pi\nu)\,e^{-0.27(j2\pi\nu)}$	$0.3(j2\pi\nu)\,e^{-0.34(j2\pi\nu)}$
	LR		0
	RL		$-0.06(j2\pi\nu)\,e^{-0.27(j2\pi\nu)}$
	RR	$0.4(j2\pi\nu)\,e^{-0.27(j2\pi\nu)}$	$0.4(j2\pi\nu)\,e^{-0.30(j2\pi\nu)}$

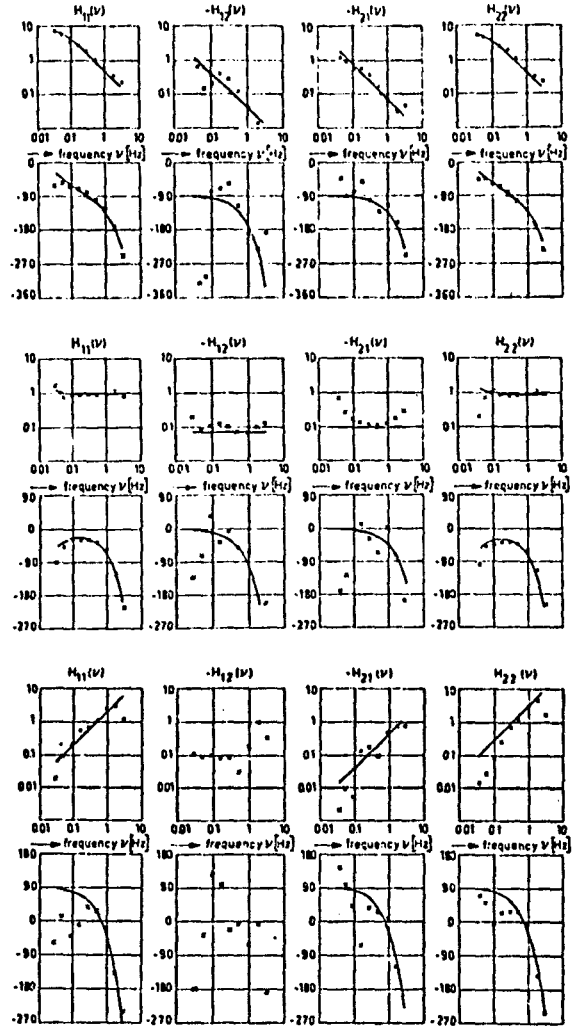


Figure 7: Raw data points (crosses) and estimated modal transfer functions (solid lines) of the averaged control behavior of the three subjects with each of the three configurations.

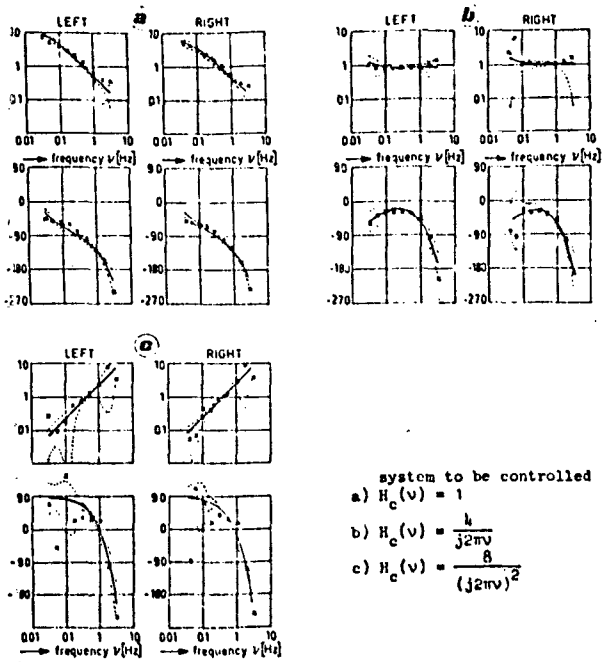


Figure 8: Raw data (crosses) and human operator model transfer function estimates (solid lines) with estimated one sigma limits (dashed lines) for the single loop experiments.

Table 5: Cross-over frequencies and phase margins.

$H_c(\nu)$	hand	single system			double system		
		ν_c[Hz]	ω_c[r/s]	φ_r[deg]	ν_c[Hz]	ω_c[r/s]	φ_r[deg]
1	L	0.53	3.3	75	0.51	3.2	75
	R	0.52	3.3	81	0.47	3.0	79
$\frac{4}{j2\pi\nu}$	L	0.57	3.6	56	0.64	4.0	46
	R	0.70	4.4	44	0.58	3.6	43
$\frac{8}{(j2\pi\nu)^2}$	L	0.51	3.2	41	0.39	2.4	43
	R	0.51	3.2	40	0.51	3.2	35

From Fig. 8 and also from Table 4 it can be seen that in the single loop task no differences occur between left and right handed tasks. The same is true for the double loop task, with an exception for the lacking left to right cross coupling in the control of the double integrator. A comparison between single loop control and double loop control shows that in the double loop experiment the transfer functions $H_{11}(\nu)$ and $H_{22}(\nu)$ show a slight increase in delay time in relation to the transfer functions in the single loop experiments. The cross-over frequencies and phase margins from Table 5 do not indicate that there are left- right differences or single loop- double loop differences. The performance measure σ_e/σ_i in Table 3 does not differ significantly between left and right or between single loop and double loop task in controlling a single or double integrator. Only in the case of controlling a proportional system the error score for a lefthanded single loop task is lower than for the right handed task or for the double loop task. This may be related to the higher value of the standard deviation σ_i of the forcing function u(t) for that case.

6 *Discussion.*

Up till now only a few investigators have been involved in identifying human operator transfer functions in multiple loop control tasks. Consistent estimation procedures with an unpredictable input and with human operator transfer functions involving a delay time were applied by Levison and Elkind [7, 8]. In [7] Levison describes a number of experiments with one control and one display. The variables of the two loops corresponded to the x- and y-direction of control and display. In the homogeneous control situation (identical systems: double integrator; equal forcing function bandwidths: 1.5, 2.5 and 3.5 rad/sec) only small differences were found between x and y and between single axis and double axis control. In [8] experiments have been described with separate displays and controls (left and right). This study was aimed at investigating the influence of display separation on human operator performance (peripheral vision, scanning). Experiments which can be compared with the present study were those with a display separation of 0.8° and controlled elements consisting of a single integrator. The human operator transfer function was modelled with four parameters: a static gain, one pole, one zero and a time delay. The main differences with reference to a one axis (foveal) control task were a slight decrease in gain (3dB) and an increase in the time delay from 0.12 to 0.15 sec.

In the present study a noticeable decrease in gain could not be found, but there was an increase in time delay of 0.03 sec. In this case a model with three parameters was chosen. A four parameter model could not be identified reliably because of the more complex overall model which included two cross-coupling terms. Even in the three parameter transfer function models the reliability will probably be lower than in the estimates of Levison and Elkind. Taking into account these differences it can be stated that the results for the transfer functions $H_{11}(\nu)$ and $H_{22}(\nu)$ do not contradict those of Levison and Elkind.

It is interesting to note that linear cross-couplings have indeed been found. These cross-couplings may have a visual origin; the moving left hand spot may influence the perception of the distance to the zero reference for the right hand spot and vice versa. The negative sign then would mean that the bias in the observation is such that the subject perceives the reference as shifted in the direction of the dot from the other control task. If, on the other hand, the cross-coupling has its origin in the motor system, it means that the hands tend to move in an opposite direction. A third possibility is that it results from central processes. Even a combination of these factors may be responsible.

Why there was no cross-coupling from left hand display to right hand control in the case of controlling two double integrators remains an intriguing question.

7 *Conclusions and Recommendations*

The study of human operator control behavior is primarily a problem of identification of a non-stationary system with a noise source in a closed loop. Therefore conclusions from this study will partly refer to the system identification problem and partly to the human operator control behavior in the task considered.

With respect to the system identification problem the following statements can be made:

- The *identification method* consisting of a *parameter estimation* procedure based on Fourier coefficients is found to be useful for this type of problem.
- The *decoupling procedure* enables the identification of *complex models* consisting of more than one describing function (cross-couplings, extra feedback element, feedforward element).
- For identification of complex models with *more than one input* it is necessary to introduce *more than one test signal* in order to identify the transfer functions from the inputs to the output(s) of the model to be identified. These test signals must be statistically independent.
- Because in the tasks observed the human control behavior can be considered *stationary only for a relatively short period* (fatigue, boredom) it is necessary to *average over a number of experiments* to obtain a reasonable signal to noise ratio in tasks involving complex human operator models.
- *Averaging* should be applied on the Fourier coefficients, i.e. *before decoupling and parameters estimation* in order to obtain a reliable choice for the structure of the describing function, and to decrease possible biases resulting from the finite observation time.
- The method applied, like most of the methods based on the use of periodic forcing functions and on estimation of Fourier coefficients, yields *consistent estimates in closed loop systems*. This means that for an infinite observation time bias and standard deviation of the estimator will be zero. The estimator is only *asymptotically unbiased* in closed loop systems.
- For a *finite observation time bias and standard deviation* of a describing function can be computed as a function of the injected noise in the system to be identified and of the dynamics of the systems in the *closed loop*.
- With *increasing observation time*, i.e. with increasing signal to noise ratio of the Fourier coefficient estimates the *bias decreases rapidly* to zero in comparison with the standard deviation.
- From the viewpoint of *system identification* a *forcing function* consisting of a sum of sinusoids should be chosen in such a way that the sinusoids should be located at those *frequencies* where a good *sensitivity* exists *to the model parameters* to be estimated.
- The *amplitudes* of the forcing function components should be chosen such that the *signal to noise ratio* in the *input* of the transfer function to be identified is about equal for all components.
- From the viewpoint of *human operator* control behavior study the forcing function should be such that the human operator is *not able to identify separate periodic components*, i.e. no single dominant components must occur in the human operator input.
- In general it can be stated that by taking into account the just mentioned recommendations and by making use of the information obtained in the present studies *a better choice for the forcing functions* should be possible for further investigations in these types of control task.

With respect to the results obtained in the double loop task the following remarks can be made:

- In controlling *two identical loops* with separate (f ly observable) displays and control *no significant differences* co found between *right* hand *and left* hand performance.
- The results obtained from execution of a *double task* l y differ from those obtained in a *single task*, only a slight increas human operator time delay (0.03 sec) could be observed.
- A *parallel model* gives a good description of the human operator control behavior in this control task. It is interesting to note the existence of *linear crosscouplings* between the two mutually independent tasks.
- The *cross-coupling terms* can be described with the *same dynamics* as the terms in the *main loops*, however, they have a *negative* sign.
- The *effect* of the cross-coupling terms is *small* in comparison with the terms in the main loops. For practical applications they can be neglected. This means that the crossover model can be applied for each of the loops separately.
- The *magnitude of the remnant* does *not differ* between *single* loop tasks and *double* loop tasks as is indicated by the standard deviations and the relative noise contents of the signals in the control loops. This means that *task interference* is only demonstrated in the linear cross-coupling terms.
- From the present data it is *not possible* to locate the *origin* (visual system, central system or motor system) of the *task interference* described by the cross-coupling terms.

8 *References*

[1] Van Lunteren, A. and H.G. Stassen: Parameter estimation in linear models of the human operator in a closed loop with application of deterministic test signals, Proc. of the 9th Annual Conf. on Manual Control, MIT, (1973), pp. 289-297.

[2] Stassen, H.G. et al.: Progress Report January 1973 until July 1976 of the Man-Machine Systems Group, Delft, Report WTHD 95, (1977), pp. 19-65.

[3] Chernikoff, R.; J.W. Duey and F.V. Taylor: Two dimensional tracking with identical and different control dynamics in each coordinate, Journal of Experimental Psychology, Vol. 60, No 5, (1960), pp. 318-322.

[4] Duey, J.W. and R. Chernikoff: The use of quickening in one coordinate of a two-dimensional tracking system, IEEE Trans. on HFE, Vol. HFE-1, (1960), pp. 21-24.

[5] Chernikoff, R. and M. Le May: Effect of various display-control configurations on tracking with identical and different coordinate dynamics, Journal of Exp. Psych.,Vol. 66, No 1, (1963),pp. 95-99.

[6] Todosiev, E.P.; R.E. Rose and L.G.Summers: Human performance in single- and two-axis tracking systems, Proc. 2nd Annual Conf. on Manual Control, NASA SP-128, (1966), pp. 143-158.

[7] Levison, W.H.: Two-dimensional manual control systems, Proc. 2nd Annual Conf. on Manual Control, NASA SP-128, (1966), pp. 159-180.

[8] Levison, W.H. and J.I. Elkind: Two-dimensional manual control systems with separate displays, Proc. 3rd Annual Conf. on Manual Control, NASA SP-144, (1967), pp. 29-42.
also: IEEE Trans. on Human Factors in Electronics, Vol. HFE-8, No 3, (1967), pp. 202-209.

[9] Van den Bos, A.: Estimation of parameters of linear systems using periodic test signals, Dr. Thesis, Delftse Universitaire Pers, (1974), 136 p.

[10] Van Lunteren, A.: Systeemidentificatie en parameterschatting in open en gesloten ketens, (System identification and parameter estimation in open and closed loops), Delft, Report N-114, (1976), pp. 85-96.

[11] Abramowitz, M. and I.A. Stegun: Handbook of mathematical functions, Dover, (1965), pp. 238-243.

[12] Jenkins, G.M. and D.G. Watts: Spectral Analysis and its applications, Holden Day, (1969), pp. 363-411.

[13] Stassen, H.G. et al. : Progress Report January 1970 until January 1973 of the Man-Machine Systems Group, Delft, Report WTHD 55, (1973), pp. 9-23.

N79-17482

AN APPROACH TO THE MULTI-AXIS PROBLEM IN MANUAL CONTROL

By Captain Walter W. Harrington

AFFDL/FGD
Wright-Patterson AFB

SUMMARY

The multi-axis control problem is addressed within the context of the optimal pilot model. The problem is developed to provide efficient adaptation of the optimal pilot model to complex aircraft systems and real world, multi-axis tasks. This is accomplished by establishing separability of the longitudinal and lateral control problems subject to the constraints of multi-axis attention and control allocation. Control solution adaptation to the constrained single axis attention allocations is provided by an optimal control frequency response algorithm. An algorithm is developed to solve the multi-axis control problem. The algorithm is then applied to an attitude hold task for a bare airframe fighter aircraft case with interesting multi-axis properties.

58

INTRODUCTION

Applications of the optimal pilot model [3-20], as well as other human operator models, have generally been limited to single axis control tasks. Methods to predict important optimal pilot model parameters, such as attention allocation [32,31] and control frequency response [34], for complex aircraft systems and tasks are recent developments. In addition, automation of the control frequency response [31] and signal to noise ratio optimal pilot model iteration loops has been accomplished only in the last few years.

This paper presents a method to solve the multi-axis control problem which is suitable for complex aircraft systems and tasks. The method takes advantage of recent developments [30,31,32,34] to be fully predictive. The method furthermore takes advantage of conventional separability assumptions for the longitudinal and lateral axis systems to provide efficient adaptation of the optimal pilot model to multi-axis tasks.

SYMBOLS

a	Threshold
A_a	Augmented open loop dynamics matrix ($n_s \times n_s$)
A_p	Augmented open loop dynamics matrix containing control filter ($n_s \times n_s$)
$\bar{A}$	Augmented closed loop dynamics matrix ($n_s \times n_s$)
B_a	Augmented control distribution matrix ($n_s \times n_c$)
C_a	Augmented measurement distribution matrix ($n_m \times n_s$)
E	Expected value
E_a	Augmented disturbance distribution matrix ($n_s \times n_d$)
f	Fraction of attention
f_{TOT}	Total attention to task
F_a	Augmented feedback matrix ($n_c \times n_s$)
g_f	Gradient of total cost with respect to fraction of attention
G	Kalman filter gain matrix
J	Control cost
J_o	Scanning cost
L	Effective feedback matrix to unaugmented state system ($n_c \times n_x$)
n_c	Number of controls
n_d	Number of disturbances
n_m	Number of measurements

n_s Number of states in augmented system

n_x Number of states in unaugmented system

P_a Riccati control gain matrix for augmented system ($n_s \times n_s$)

P_m Motor noise to signal ratio

P_y Measurement noise to signal ratio

Q Measurement penalty matrix ($n_m \times n_m$)

R Control rate penalty matrix ($n_c \times n_c$)

tr Trace

T Gaussian random input describing function approximation for a threshold nonlinear

$\underline{u}$ Pilot's control input, a vector of dimension n_c

$\underline{V}_m$ Autocovariance of motor noise, a vector of dimension n_c

$\underline{V}_y$ Autocovariance of measurement noise, a vector of dimension n_m

$\underline{w}$ A disturbance vector of Gaussian white noise, a vector of dimension n_d

$\underline{x}_a$ State of the augmented system, a vector of dimension n_s

X_a State covariance matrix of the augmented system ($n_s \times n_s$)

$\underline{y}$ A vector of measurements available to the pilot, of dimension n_m

Σ Riccati filter covariance matrix

τ Pure time delay

ω_c Control cutoff frequency

ω_N Neuromuscular cutoff frequency

Ω_c Control filter matrix

Subscripts

a Augmented

p Perceived

Superscripts

T Transpose

* Optimal

^ Estimated parameter

OPTIMAL PILOT MODEL

The optimal pilot model concept, developed by Kleinman, Baron, and Levison [3-20], has demonstrated success in modeling complex, time varying control tasks. The optimal pilot model is a mathematical construct designed to synthesize pilot control performance and behavior. The model is based on the assumption that the human operator will control a dynamic, stochastic system optimally subject to his inherent limitations. These limitations are considered to be

1. A time delay, representing cognitive, visual central processing, and neuromotor delays.

2. "Remnant" signals, divided into an observation noise to represent signal degradation due to work load, scanning effects, and signal thresholds, and a motor noise to represent random errors in executing the intended control.

3. A "neuromuscular lag" to represent neuromuscular dynamics.

The control commands are synthesized by a continuous linear equalization network which contains a full state optimal filter (Kalman) filter), a full state optimal predictor, and a full state optimal feedback control law. The control law is derived for an augmented state system which results from introducing the neuromuscular lag by means of a control rate penalty. The structure of the model results from a suboptimal solution to a control problem involving a time delay and observation noise. The model is shown in Figure 1.

The mathematical algorithm of the optimal pilot model is developed from the following control problem:

Given the quadratic cost functional of the form

$$J = 1/2 \int_0^\infty E\{\underline{y}^T(t)\, Q\, \underline{y}(t) + \underline{u}_a^T(t)\, R\, u_a(t)\}\, dt \qquad (1)$$

subject to the constraints

$$\dot{\underline{x}}_a(t) = A_a \underline{x}_a(t) + B_a \underline{u}_a(t) + E_a \underline{w}_a(t) \qquad (2)$$

$$\underline{y}(t) = C_a \underline{x}_a(t - \tau) + \underline{v}_y(t - \tau), \qquad (3)$$

determine the non-anticipative feedback control $\underline{u}_a^*(t)$ which minimises the cost functional.

Optimal Control Law

The optimal pilot model equalization network contains a full-state optimal feedback control law. The pilot model determines the feedback control $\underline{u}_a^*(t)$ which minimises the cost functional

$$J = \frac{1}{2} \int_0^\infty E\{\underline{y}^T(t) Q \underline{y}(t) + \underline{u}_a^T(t) R\, \underline{u}_a(t)\} dt \qquad (1)$$

subject to the constraints

$$\dot{\underline{x}}_a(t) = A_a \underline{x}_a(t) + B_a \underline{u}_a(t) \qquad (2a)$$

$$\underline{y}(t) = C_a \underline{x}(t) \qquad (3a)$$

by the solution of the steady state Riccati control matrix equation

$$A_a^T P_a + P_a A_a + C_a^T Q C_a - P_a B_a R^{-1} B_a^T P_a = 0 \qquad (4)$$

Figure 1 Structure of Optimal Pilot Model

where

$$\underline{u}_a^*(t) = -R^{-1}B_a^T P_a \underline{x}_a(t).$$

The control law therefore requires the specification of the measurement penalty matrix Q and the control rate penalty matrix R. The measurement penalty matrix Q is specified to provide rms minimization of the measured quantities. The control rate penalty matrix R is related to control frequency response in the following section.

Control Frequency Response

Pilot control frequency is regulated in the optimal pilot model by a first order filter matrix which processes the augmented control signals such that

$$\underline{u}_a^*(t) = \dot{\underline{u}}(t) = -\Omega_c L\underline{x}(t) - \Omega_c \underline{u}(t) \tag{5}$$

It can be shown [31,34] that the filter matrix, Ω_c, is given by

$$\Omega_c = R^{-1} P_f \tag{6}$$

where

$$P_a = \left[\begin{array}{c|c} P & P_e \\ \hline P_e & P_f \end{array}\right] \tag{7}$$

is obtained from Equation (4).

The diagonal elements ω_{c_i}, i=1, ..., n_c, of Ω_c represent the first order cutoff frequencies of the control inputs u_i, i=1, ..., n_c. The cutoff frequencies are constrained such that

$$) \leq \omega_{c_i} \leq \omega_{N_i}, \quad i=1, \ldots, n_c \tag{8}$$

where $\omega_{c_i} \geq 0$ by definition and ω_{N_i} is the pilot's neuromuscular frequency limit for the ith control input.

Iterative techniques have been developed [31] by which the control solution can be regulated so that the cutoff frequencies ω_{c_i}, i=1, ..., n_c, attain any desired set of values, subject to the constraints of Equation (8). These values can be predicted by the model through the optimal control frequency response algorithm [34]. This algorithm solves a complex, nonlinear optimization problem to determine the set of cutoff frequencies which minimize the cost function given by Equation (1) subject to the constraints of Equation (8). The algorithm normally forms an outer loop about the basic pilot model, requiring actual cost evaluation and optimization.

Full State Estimation

The optimal pilot model equalization network generates a full state estimate of the aircraft system based on noisy, delayed observations of the system. This estimate is the output of a Kalman filter. The filter gains are determined from the covariance matrix Σ which is the solution to the steady state Riccati filter equation

$$A_p\Sigma + \Sigma A_p^T + W - \Sigma C_a^T V_y^{-1} C_a \Sigma = 0 \tag{9}$$

where

$$W = E_a V_d E_a^T \tag{10}$$

The matrix V_d is the variance matrix of state disturbances of which the autocovariances of the motor noise, $\underline{V}_m$, are elements. Also,

$$\underline{V}_y = \text{Diagonal } (V_y) \tag{11}$$

where $\underline{V}_y$ is the autocovariance of the measurement noise. The filter therefore requires specification of the autocovariance of the measurement noise $\underline{V}_y$ associated with the measurements $\underline{y}(t)$ and the autocovariance of the motor noise $\underline{V}_m$ associated with the controls $\underline{u}(t)$.

Pilot Model Measurements

The optimal pilot model is well suited for realistic synthesis of human operator information processing. The pilot model can observe those instruments or quantities which the pilot would observe to perform the flight task. Furthermore, the pilot model contains algorithms for amplitude and rate information processing, attention allocation, and signal perception.

Measurement Noise

The optimal pilot model contains measurement noise to represent signal degradation due to work load, attention allocation, and signal thresholds. It is assumed that the covariance V_{y_i} of each measurement noise $v_{y_i}(t)$, $i=1, \ldots, n_m$, is given by

$$V_{y_i} = \frac{\pi P_y}{f_i T^2} E(y_i^2(t)), \quad i=1, \ldots, n_m \tag{12}$$

where P_y is the full attention noise to signal ratio,

$$P_y = .01, \tag{13}$$

f_i is the fraction of attention allocated to $y_i(t)$, subject to the constraints

$$\sum_{i=1}^{n_m} f_i = f_{TOT} \leq 1 \tag{14}$$

$$f_i \geq 0 \tag{15}$$

and T is a Gaussian input describing function approximation for a threshold of a_i.

$$T = \operatorname{erfc} \frac{a_i}{\sqrt{2}\,\sigma_i} \tag{16}$$

where $\sigma_i = \sqrt{E(y_i^2(t))}$. Iteration of the covariance calculations is required since the autocovariance of the motor noise is a function of system performance. A method for predicting the attention allocation is presented in the following section.

Attention Allocation

An optimal attention allocation algorithm has been developed by Kleinman [32] to solve the optimal attention allocation problem. The basic hypothesis of optimal attention allocation is that the human pilot will adapt his attention allocation so as to minimize the cost functional given by Equation (1) subject to the constraints given by Equations (14) and (15). The optimization process developed by Kleinman is carried out numerically via a gradient algorithm which determines an unconstrained gradient vector, g_f. An optimization algorithm then solves the constrained optimization problem.

Motor Noise

The optimal pilot model contains motor noise to represent random errors in the execution of the intended control as well as control signal degradation due to work load and attention allocation. The autocovariance or the motor noise is given by [31]

$$V_{m_i} = \frac{\pi P_m}{f_{TOT}} E(u_i^2(t)), \quad i=1, \ldots, n_c \tag{17}$$

where P_m is the full attention allocation noise to signal ratio,

$$P_m = .003 \tag{18}$$

and f_{TOT} is the total attention to the task. The autocovariance of the motor noise completes the specification of quantities required by the Kalman filter.

Full State Prediction

The optimal pilot model equalization network contains an optimal, full state predictor which updates the delayed full state estimate generated by the Kalman filter. It is required by the predictor to specify the pure time delay τ by which the observations $\underline{y}(t)$ are delayed.

Pilot Model Performance Prediction

The optimal pilot model synthesizes pilot control performance by the generation of piecewise constant aircraft system statistics. The state covariance matrix X arises from the sum of the covariance of the filter estimate, the covariance of the estimation error, and the covariance of the predictor error [8]. Thus,

$$X = E(\underline{x}_a(t)\,\underline{x}_a^T(t)) \tag{19}$$
$$= \int_0^\tau e^{\bar{A}\sigma} e^{A_p\tau}\,\Sigma\, C_a V_y^{-1} C_a L e^{A_p^T\tau} e^{\bar{A}\sigma}\,d\sigma$$
$$+ e^{A_p\tau}\,\Sigma\, e^{A_p\tau} + \int_0^\tau e^{A_p\sigma} W e^{A_p^T\sigma}\,d\sigma$$

The standard deviation of the aircraft states is then given by

$$\sigma_{x_i} = \sqrt{E(x_i^2(t))} = \sqrt{(X)_{ii}} \tag{20}$$

The standard deviation of the measurements is given by

$$\sigma_{y_i} = \sqrt{(C_a X C_a^T)_{ii}} \tag{21}$$

In addition to the rms statistics, the optimal pilot model has predicted the control frequency response and attention allocation. Pilot ratings [29] can also be predicted.

APPLICATION OF THE OPTIMAL PILOT MODEL TO COMPLEX AIRCRAFT SYSTEMS AND MULTI-AXIS TASKS

Typical aircraft systems contain many subsystems, including a control feel system, control actuators, the bare airframe, a stability augmentation system, and instrument and visual display systems. Since the optimal pilot model requires a linear representation of the aircraft system dynamics, these subsystems must be represented by sets of first order, linearized, piecewise constant coefficient, differential equations of motion. Statistical linearization techniques can be applied to nonlinearities. Iteration is then required to adapt these coefficients to the resulting system performance.

The resulting state model of the aircraft system dynamics can be quite large, requiring excessive time and core requirements for computer implementation of the optimal pilot model. However, several considerations can result in significant reduction of these requirements.

Decoupling of the Optimal Control Law

The linearized representation of conventional aircraft systems results in decoupled sets of longitudinal and lateral equations. If the pilot does not intentionally couple the longitudinal and lateral axis systems, i.e., Q and R are block diagonal in a manner similar to the decoupling of the system equations, the control problem formed by seeking the feedback control $\underline{u}_a^*(t)$, which minimizes J subject to the constraints of Equations (2a) and (3a), decouples into two separate control problems. The measurement penalty matrix Q is normally specified to be diagonal, which of course satisfies the constraint of being block diagonal. In addition, the control rate penalty matrix R is normally specified to be diagonal for the control frequency response regulation algorithm [31]. Thus, the control law can be determined for each axis system for multi-axis tasks.

Control Frequency Response

Control frequency response, as specified by the optimal control frequency response algorithm [31,34], is dependent upon predicted system performance. However, the assumptions of the previous section imply that the control filter matrix, Ω_c, given by Equation (6) is block diagonal. Thus, if the covariance propagation problem can be constrained for each axis system, the optimal control frequency response problem decouples into two separate control problems.

Full State Estimation

The separability of the state estimation problem is dependent upon the specification of the state disturbances and measurement noise. The state disturbances, including the motor noise, are usually model by independent Gaussian white noise sources, which are separable. It is proposed that f_{TOT} in Equation (17)

$$\underline{V}_{m_i} = \frac{\pi P_m}{f_{TOT}}\, E(u_i^2(t)),\ i=1,\ \ldots,\ n_c \tag{17}$$

be represented by the total attention allocation for each axis system. The uncertainty of control is then related to the attention to control for each axis system.

The measurement noises are decoupled for each axis system with the exception of dependence upon attention allocation. Thus, the state estimation problem is coupled by the attention allocation algorithm.

Attention Allocation

The optimal attention allocation algorithm developed by Kleinman [32] has two major components: the gradient algorithm which determines an unconstrained gradient vector, and the optimization algorithm which solves the constrained optimization problem. The gradient algorithm is decoupled for each axis system. However, the optimization problem must be solved for the entire aircraft system.

If the total attention allocation is constrained for each axis system for each attention optimization iteration, the state estimation problem can be decoupled. The optimal attention allocation gradient algorithm can then be applied separately to each axis system. The optimal attention allocation optimization algorithm is then applied to the entire system. Note that the optimization algorithm should not be applied to each axis system, since this will only identify local minima, as well as incur additional computation cost.

Covariance Propagation

The remaining covariance propagation problems are decoupled into the longitudinal and lateral axis systems. If the attention allocation is constrained for each axis system for each attention optimization iteration as described in the previous section, the covariance propagation problem is also constrained, as required for decoupling of the optimal control frequency response problem. Note that the attention allocation optimization algorithm now forms an outer iteration loop about the optimal control frequency response algorithm. This is satisfactory since the two algorithms are mutually adaptive to each other.

Computation Requirements

The multi-axis control problem can be decoupled into the longitudinal and lateral control problems with the exception of attention allocation optimization. The separate solution of longitudinal and lateral control problems will require much less computation time due to the reduction in matrix size. Furthermore, the same core can be used for both solutions. Thus, significant time and core savings can be realized by this approach.

Flow charts of the computations required by this approach are shown in Figures 2 and 3. The computations are structured so as not to interfere with the utility of the model for single axis tasks. The multi-axis attention optimization, as well as storage of required pilot model parameters, is therefore performed by a separate subroutine.

APPLICATION TO AN ATTITUDE HOLD TASK

The multi-axis control task scheme is applied to an attitude hold task for a bare airframe fighter aircraft case. An attitude hold task for a bare airframe provides a simple example which can be easily duplicated for further research in this area. The bare airframe fighter aircraft case selected has unstable longitudinal dynamics and more conventional lateral dynamics, which will be useful for the illustration of basic properties of the multi-axis control task scheme.

Aircraft System

The fighter aircraft case involves straight and level flight at an altitude of 3,048 meters (10,000 feet) at an airspeed of 262 meters/second (862 feet/second). The airframe dynamics for this case are modeled by standard, linearized, primed, longitudinal and lateral body axis equations of motion [31]. The stability derivatives and other parameters pertinent to this case are presented in Table 1. For simplicity, as well as to accentuate the aircraft dynamics, no models are included for the stability augmentation system or control feel system. The airframe is disturbed by turbulence with gust intensities of 5 feet/second. MIL SPEC 8785B turbulence is provided, as modeled by Heath [24].

Controls include a side stick for pitch and roll and pedals for yaw. The neuromuscular frequency response limits are estimated to be 10, 10, and 5 radians/second for the pitch stick, roll stick, and pedals, respectively [25,26]. The multi-dimensional control for the lateral axis system will be useful to demonstrate the utility of the multi-axis control task algorithms.

Measured quantities include pitch angle, vertical velocity, roll angle, and yaw angle. It is assumed that the rates of these quantities are simultaneously available. The perceptual thresholds for the measured quantities are presented in Table 2.

Task

The task is defined to be precision attitude hold in turbulence. Control action is required as soon as a deviation in measured quantities is noticed, and deviations must be maintained within specified limits. A total of 80% attention is allocated to this task.

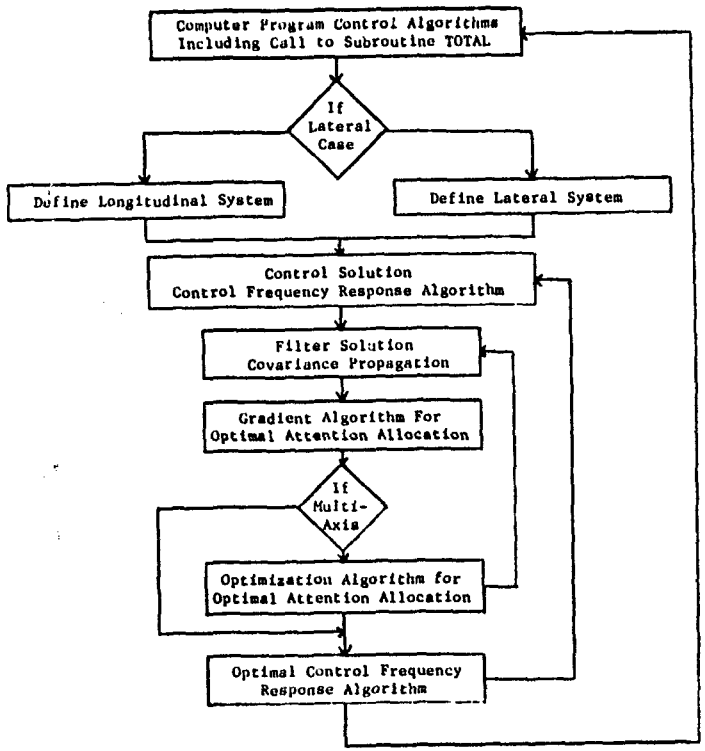


Figure 2. Main Program Structure for Multi-Axis Problem

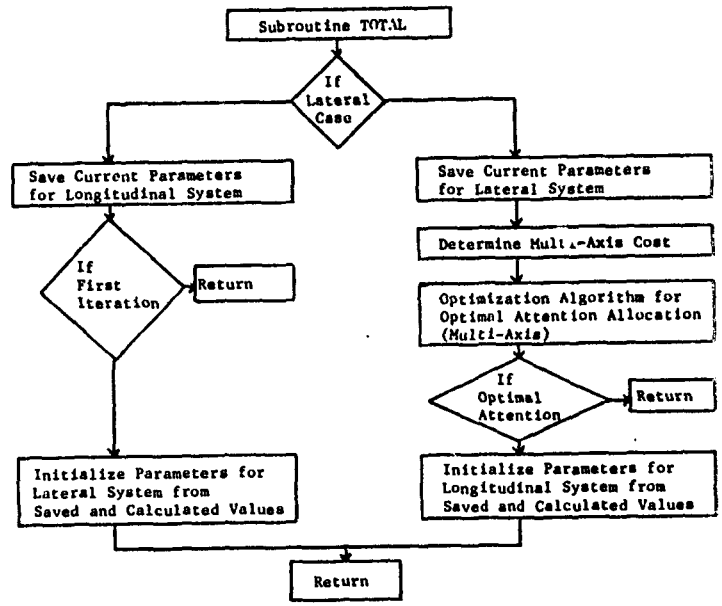


Figure 3. Subroutine TOTAL Structure for Multi-Axis Problem

PARAMETER	LONGITUDINAL AXIS SYSTEM	LATERAL AXIS SYSTEM
Unprimed Stability Derivatives	X_u = -.01578 X_w = .02144 $X_{\delta e}$ = 20.6 M_u = -.001169 M_w = .005766 M_q = -.6528 $M_{\delta e}$ = -28.2 $M_{\dot{w}}$ = -.0003094 Z_u = -.06086 Z_w = -1.692 $Z_{\delta e}$ = -157.3	Y_v = -.2932 $Y_{\delta a}$ = 14.15 $Y_{\delta r}$ = 50.72 L_β = -70.95 L_p = -2.551 L_r = -.06964 $L_{\delta a}$ = -62.92 $L_{\delta r}$ = 17.01 N_β = 11.51 N_p = -.01704 N_r = -.2266 $N_{\delta a}$ = -2.261 $N_{\delta r}$ = -9.057
Moments		I_{xx} = 9017.5 I_{zz} = 58104. I_{xz} = 144.93
Wing Span	30.	30.
Total Velocity	862.0	862.0
Flight Path Angle	0.	0.
Angle of Attack	1.5285	1.5285
Altitude	10000.	10000.
Units	Feet, Seconds, Radians	Feet, Seconds, Radians

TABLE 1. FIGHTER AIRCRAFT STABILITY DERIVATIVES AND OTHER PARAMETERS

MEASURED QUANTITY	UNITS	PERCEPTUAL THRESHOLD	INDIFFERENCE THRESHOLD	MAXIMUM ALLOWABLE DEVIATION	STATE PENALTY
Pitch Angle	deg.	.8	.8	10.	.01
Pitch Rate	deg./sec.	1.6	1.6	32.	.001
Vertical Velocity	ft./sec.	.27	.27	10.	.01
Vertical Acceleration	ft./sec.^{2}	.54	.54	32.	.001
Roll Angle	deg.	.9	.9	10.	.01
Roll Rate	deg./sec.	1.8	1.8	32.	.001
Yaw Angle	deg.	.9	.9	10.	.01
Yaw Rate	deg./sec.	1.8	1.8	32.	.001

Table 2. Measured Quantities and Related Parameters

Immediate control action implies that the indifference thresholds are the same as the perceptual thresholds, as presented in Table 2. The maximum allowable deviations are also presented in this table. The optimal pilot model state penalties are derived by the conventional rule

$$Q_{ii} = \left(\frac{1}{\text{MAX}|y_i(t)|}\right)^2 \qquad (31)$$

and are presented in Table 2.

Model Predictions

The multi-axis control task scheme is implemented on the AFFDL optimal pilot model computer program FGDPILOT [31], which is designed for complex aircraft system and simulation analysis. The computer program is completely automated, and requires essentially only those parameters presented in this section. The program contains nominal values for all the basic optimal pilot model parameters. Program predictions include optimal control frequency response, optimal attention allocation [32], control and aircraft system performance, and pilot rating [29]. Predictions for this case are presented in Figure 4.

The predicted attention allocation for this case is presented in Table 3. This attention allocation is shown in comparison with the single axis results obtained in reference [34], where 40% attention was allocated to each axis system. The multi-axis solution indicates a large shift in attention allocation to the unstable longitudinal axis system, as would be expected.

OBSERVED QUANTITY	SINGLE AXIS ATTENTION ALLOCATION	TOTAL SINGLE AXIS ATTENTION ALLOCATION	MULTI-AXIS ATTENTION ALLOCATION	TOTAL MULTI-AXIS ATTENTION ALLOCATION
Pitch Angle	.031		.200	
Pitch Rate	.031	.400	.200	.697
Vertical Velocity	.369		.497	
Vertical Accel	.369		.497	
Roll Angle	.304		.086	
Roll Rate	.304	.400	.086	.103
Yaw Angle	.096		.017	
Yaw Rate	.096		.017	

Table 3. Predicted Attention Allocation


```
CASE DESCRIPTION

RARE AIRFRAME DYNAMICS, M=.8
6-DOF PRECISION ATTITUDE HOLD TASK WITH PITCH STICK, ROLL STICK AND PEDALS
LONGITUDINAL DYNAMICS

ALTITUDE = 1.000000E+04 FEET
NOMINAL AIRSPEED = 8.616933E+02 FEET/SECOND
NOMINAL ANGLE OF ATTACK = 1.528500E+00 DEGREES
TURBULENCE CONDITIONS
     SIG U GUST = 5.000000E+00 FEET/SECOND
     SIG V GUST = 5.000000E+00 FEET/SECOND
     SIG W GUST = 5.000000E+00 FEET/SECOND

AIRFRAME EIGENVALUES
(-3.6681283)+J(0.)
(1.0855377)+J(0.)
(-.22298678E-01)+J(.13527864)
(-.22298678E-01)+J(-.13527864)

                    PILOT MODEL PARAMETERS

TOTAL ATTENTION ALLOCATION TO PRIMARY FLIGHT TASK (ATTN) = .696898E+00
PURE TIME DELAY (TAU) = .200000E+00 SECONDS
MOTOR NOISE TO SIGNAL RATIO (PY) = 1.000000E-02
TOTAL COST (COST) = 1.285988E+00
SCANNING COST (SCOST) = 8.281187E-01
```

Figure 4. Optimal Pilot Model Performance Predictions for Multi-Axis F-16 Bare Airframe Attitude Hold Task (1 of 4)

CONTROLS

CONTROLLER	NEUROMUSCULAR FREQUENCY	OPTIMAL CONTROL FREQUENCY	CONTROL RATE PENALTY (RINV)	CONTROL FORCE STD DEVIATION	CONTROL SURFACE STD DEVIATION
PITCH STICK	10.000000	10.317287	9.9196789	.21455517	.21455517

OBSERVED QUANTITIES

OBSERVED QUANTITY	UNITS	TASK VECTOR	PERCEPTUAL THRESHOLD	INDIFFERENCE THRESHOLD	ATTENTION ALLOCATION	STANDARD DEVIATION
PITCH ANGLE	DEG	.10000000E-01	.80000000	.80000000	.19989588	3 5283864
PITCH RATE	DEG/SEC	.10000000E-02	1.6000000	1.6000000	.19989588	1.4511777
VERTICAL VELOCITY	FT/SEC	.10000000E-01	.27000000	.27000000	.49700244	7.1923670
VERTICAL ACCEL	FT/SEC**2	.10000000E-02	.54000000	.54000000	.49700244	21.733312

Figure 4. Optimal Pilot Model Performance Predictions for Multi-Axis F-16 Bare Airframe Attitude Hold Task (2 of 4)

CASE DESCRIPTION

BARE AIRFRAME DYNAMICS, M = .8
6-DOF PRECISION ATTITUDE HOLD TASK WITH PITCH STICK, ROLL STICK AND PEDALS
LATERAL DYNAMICS

ALTITUDE = 1.000000E+04 FEET
NOMINAL AIRSPEED = 8.616933E+02 FEET/SECOND
NOMINAL ANGLE OF ATTACK = 1.528500E+00 DEGREES
TURBULENCE CONDITIONS
SIG U GUST = 5.000000E+00 FEET/SECOND
SIG V GUST = 5.000000E+00 FEET/SECOND
SIG W GUST = 5.000000E+00 FEET/SECOND

AIRFRAME EIGENVALUES
(0.)+J(0.)
(-2.5341997)+J(0.)
(-.1876088E-01)+J(3.6290849)
(-.25919919)+J(-3.6290849)

PILOT MODEL PARAMETERS

TOTAL ATTENTION ALLOCATION TO PRIMARY FLIGHT TASK (ATTN) = .103102E+00
PURE TIME DELAY (TAU) = .200000E+00 SECONDS
MOTOR NOISE TO SIGNAL RATIO (PM) = 3.000000E-03
MEASUREMENT NOISE TO SIGNAL RATIO (PY) = 1.000000E-02
TOTAL COST (COST) = 3.528117E-01
SCANNING COST (SCOST) = 1.290835E-01

Figure 4. Optimal Pilot Model Performance Predictions for Multi-Axis F-16 Bare Airframe Attitude Hold Task (3 of 4)

The predicted control frequency responses for this case are presented in Table 4. These results are shown in comparison with the results obtained in reference [34]. The maximum physically attainable frequency response is again predicted for the unstable longitudinal axis system. But, a substantial reduction in control frequency response is predicted for the lateral axis system. This reduction is most consistent with the reduction in attention allocation to this axis system. Note that the pedal cutoff frequency of .2 radians/ second corresponds to a control cycle time of about 30 seconds. These values could not be predicted by previous methods.

CONTROLLER	NEUROMUSCULAR CUTOFF FREQUENCY	SINGLE AXIS OPTIMAL CONTROL CUTOFF FREQUENCY	MULTI-AXIS OPTIMAL CONTROL CUTOFF FREQUENCY
Pitch Stick	10.rad/sec	10.rad/sec	10.rad/sec
Roll Stick	10.rad/sec	7.6rad/sec	4.2rad/sec
Pedals	5.rad/sec	4.4rad/sec	.20rad/sec

Table 4. Predicted Control Frequency Response

The multi-axis solution also produces a substantial reduction in total quadratic cost, as shown in Table 5. This is further evidence of the adaptive capability of the methods developed herein and in reference [34].

AXIS SYSTEM	SINGLE AXIS QUADRATIC COST	MULTI-AXIS QUADRATIC COST
Longitudinal	2.51	1.29
Lateral	.18	.35
TOTAL SYSTEM	2.69	1.64

Table 5. Predicted Quadratic Cost

SUMMARY OF THE MULTI-AXIS CONTROL PROBLEM

A method has been presented to efficiently solve the multi-axis control problem which is suitable for complex aircraft systems and

Figure 4. Optimal Pilot Model Performance Predictions for Multi-Axis F-16 Bare Airframe Attitude Hold Task (4 of 4)

CONTROLS

CONTROLLER	NEUROMUSCULAR FREQUENCY	OPTIMAL CONTROL FREQUENCY	CONTROL RATE PENALTY (RMV)	CONTROL FORCE STD DEVIATION	CONTROL SURFACE STD DEVIATION
ROLL STICK	10.000000	4.1501079	11.986371	.20936254	.20936254
PEDALS	5.000000	.20265449	.45316230	.24956347	.24956347

OBSERVED QUANTITIES

OBSERVED QUANTITY	UNITS	TASK VECTOR	PERCEPTUAL THRESHOLD	INDIFFERENCE THRESHOLD	ATTENTION ALLOCATION	STANDARD DEVIATION
ROLL ANGLE	DEG	.10000000E-01	.90000000	1.8000000	.85765583E-01	4.4072659
ROLL RATE	DEG/SEC	.10000000E-02	1.8000000	1.8000000	.85765583E-01	10.R8896
YAW ANGLE	DEG	.10000000E-01	.90000000	.90000000	.17336093E-01	1.2494873
YAW RATE	DEG/SEC	.10000000E-02	1.8000000	1.8000000	.17336093E-01	2.4036528

tasks. The method has been applied to an attitude hold task for a bare airframe fighter aircraft case. This application demonstrated the method's capability to make realistic predictions for stable as well as unstable aircraft system dynamics and for scalar as well as multi-dimensional controls. The method shows great promise for complex aircraft system and simulation analysis.

REFERENCES

1. Gressang, Stone, Pollard, and Kugel: Low Visibility Landing Pilot Modeling Experiment and Data, Phase I. AFFDL TR 75-41, April 1976.

2. Gressand: Low Visibility Landing Pilot Modeling Experiments and Data, Phase II. AFFDL TR 75-57, August 1975.

3. McRuer and Krendel: Mathematical Models of Human Pilot Behavior. AGARDograph No. 188, January 1974.

4. Baron and Kleinman: The Human As An Optimal Controller and Information Processor. IEEE Trans. Man-Machine Systems, Vol. MMS-10, March 1969, pp 9 - 17.

5. Baron et al: Application of Optimal Control Theory to the Prediction of Human Performance In A Complex Task. AFFDL TR 69-81, March 1970.

6. Kleinman, Baron, and Levison: A Control Theoretic Approach to Manned-Vehicle Systems Analysis. IEEE Trans. Auto. Control, Vol. AC-16, December 1971, pp 824-832.

7. Kleinman, Baron, and Levison: An Optimal-Control Model of Human Response, Part I: Theory and Validation. Automatica, Vol. 6, 1970, pp 357-369.

8. Kleinman: Optimal Control of Linear Systems with Time-Delay and Observation Noise. IEEE Trans. Auto. Control, Vol. AC-14, October 1969, pp 524-527.

9. Graham and McRuer: Analysis of Nonlinear Control Systems. Dover Publications, New York, 1971, Chapter 6.

10. Baron and Levison: An Optimal Control Methodology for Analyzing the Effects of Display Parameters On Performance and Work Load in Manual Flight Control. IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-5, July 1975, pp 423-430.

11. Bryson and Ho: Applied Optimal Control. Ginn and Co., Waltham, Mass., 1969, Chapter 5.

12. Kleinman and Baron: Analytic Evaluation of Display Requirements for Approach to Landing. NASA CR-1952, October 1971.

13. Kleinman and Perkins: Modeling Human Performance in a Time-Varying Anti-Aircraft Tracking Loop. IEEE Trans. Auto. Control, Vol. AC-19, August 1974, pp 297-306.

14. Harvey: Application of An Optimal Control Pilot Model to Air-to-Air Combat. AFIT Thesis GA/MA/74M-1, March 1974.

15. Dillow, Picha and Anderson: Slushy Weightings for the Optimal Pilot Model. 11th Annual Conference on Manual Control, May 1975.

16. Kleinman: Computer Programs Useful in Linear Systems Studies. Systems Control, Inc., Technical Memorandum, December 1971.

17. Kleinman: On An Iterative Technique for Riccati Equation Computation. IEEE Trans. Auto. Control, Vol. AC-13, February 1968, pp 114-115.

18. Kwakernaak and Sivan: *Linear Optimal Control Systems*. John Wiley, New York, 1972, Chapter 3.

19. McRuer, Ashkenas, and Graham: Aircraft Dynamics and Automatic Control. Princeton University Press, 1973, Chapters 5 and 6.

20. Kleinman and Killingsworth: A Predictive Pilot Model for STOL Aircraft Landing. NASA CR-2374, March 1974.

21. Gressang: A Note on Solving Riccati Equations Associated with Optimal Pilot Models. AFIT/AIAA Mini Symposium on Recent Advances in Aeronautical Research Development, and Systems, WPAFB, 26 March 1975.

22. Bryson and Hall: Optimal Control and Filter Synthesis by Eigenvector Decomposition. Report, Dept. of Aero. and Astro., Stanford University, December 1971.

23. McRuer and Graham: Human Pilot Dynamics in Compensatory Systems. AFFDL TR 69-72, July 1965.

24. Heath: State Varible Model of Wind Gusts. AFFDL/FGC TM-72-12, July 1972.

25. Stark: Neurological Control Systems, Studies in Bioengineering. Plenum Press, New York, 1968.

26. Milsum: Biological Control Systems Analysis. McGraw-Hill Book Company, New York, 1966.

27. Curry, Young, Hoffman, and Kugel: A Pilot Model with Visual and Motion Cues. AIAA Visual and Motion Simulation Conference, Dayton, Ohio, April 1976.

28. Harrington: The Application of Pilot Modeling to the Study of Low Visibility Landing. Twelfth Conference on Manual Control, May 1976.

29. Hess: Prediction of Pilot Opinion Ratings Using An Optimal Pilot Model. To appear in Human Factors.

30. Curry, Hoffman, Young: Pilot Modeling for Manned Simulation. AFFDL-TR-76-124, Volume 1, December 1976.

31. Harrington: A Computer Program for the Analysis of Manned Aircraft and Simulation Systems. AFFDL TR to be published.

32. Kleinman: Solving the Optimal Attention Allocation Problem in Manual Control. IEEE Trans. Auto. Control, Vol. AC-21, No. 6, December 1976, pp 813-821.

33. Hoffman, Curry, Kleinman, Hollister, Young: Display/Control Requirements for VTOL Aircraft. ASI-TR-75-26 (NASA CR 145026), August 1975.

34. Harrington: The Optimal Control Frequency Response Problem in Manual Control. Thirteenth Conference on Manual Control, June 1977.

ERROR RATE INFORMATION IN ATTENTION ALLOCATION PILOT MODELS

W. H. Faulkner
E. D. Onstott

Northrop Corporation
Aircraft Group
Hawthorne, California

Proceedings of the
Thirteenth Annual Conference
On Manual Control
June 15-17, 1977

Massachusetts Institute of Technology
Cambridge, Massachusetts

ERROR RATE INFORMATION IN ATTENTION ALLOCATION PILOT MODELS

W. H. Faulkner
E. D. Onstott

Northrop Corporation
Aircraft Group
Hawthorne, California

ABSTRACT

The Northrop urgency decision pilot model was used in a command tracking task to compare the optimized performance of multi-axis attention allocation pilot models whose urgency functions were 1) based on tracking error alone, and 2) based on both tracking error and error rate. A matrix of system dynamics and command inputs was employed, to create both symmetric and asymmetric two-axis compensatory tracking tasks. All tasks were single loop on each axis. Analysis showed that a model that allocates control attention through nonlinear urgency functions using only error information could not achieve performance of the full model whose attention shifting algorithm included both error and error rate terms. Subsequent to this analysis, tracking performance predictions for the full model were verified by piloted flight simulation. Complete model and simulation data are presented.

INTRODUCTION

A great deal is known about the dynamics of the human pilot performing continuous linear single axis tasks. Much work has gone into developing models that match the amplitude and phase characteristics of the pilot's output at the controller, and many aspects of the internal structure of the human have been analyzed. Such models are of use in solving many human factors problems, but for the basic objective of determining the total system dynamics, it is usually sufficient to employ simple models that consist of gain K_p, lead T_L, time delay τ, and possibly lag T_I:

$$Y_p = K_p \frac{(T_L s + 1)}{(T_I s + 1)} e^{-\tau s} \qquad (1)$$

In multi-axis tasks, the human controller must behave as a time shared system element. This shifting of attention degrades the pilot's performance on each axis from what he could achieve in continuous control. This attention shifting is by no means random or regular; the pilot is quite discriminating about when he will abandon

the control of one axis to take over the control of another. This leads to a pilot attention shifting criterion that is functionally dependent on the total system variables.

Thus, any model that attempts to extend single axis theory to multi-axis tasks must contain an algorithm that determines when attention shifting takes place. The Northrop urgency decision pilot model employs for this purpose nonlinear functions, called urgency functions, of the state variables of each axis. For most tasks, the urgency function of an axis x is of the form:

$$U_x = \left| \alpha_x |\epsilon_x| + \beta_x \frac{\epsilon_x}{|\epsilon_x|} \dot{\epsilon}_x \right| \qquad (2)$$

where α and β are constants, and ϵ and $\dot{\epsilon}$ are the error and error rate respectively for the axis in question. Then the attention shifting algorithm of the model is simply this: the axis having the largest urgency function receives the corrective control attention.

The Northrop multi-axis model thus consists of the simple linear dynamics of equation (1) along with the control criterion of equation (2). The linear coefficients are chosen so that the model performs optimally in the particular task under consideration.

ERROR RATE INFORMATION IN ATTENTION ALLOCATION PILOT MODELS

The postulated form of the urgency functions in the Northrop urgency decision pilot model has provision for both error and error rate terms. It was the purpose of the present study 1) to apply the optimized complete model with urgency function error rate terms to marginally stable and unstable second order system dyramics, and 2) to examine the characteristics of the performance of the optimized incomplete model without urgency function error rate terms. In this way, the importance of the error rate terms in achieving optimum performance can be assessed by examining the control strategy the model adopts to compensate for the lack of error rate information in its attention shifting algorithm.

The postulated form of the urgency functions has led to correct predictions in VTOL hover, attitude stabilization in turbulence, and air-to-air target tracking analyses as reported in References 1, 2, and 3. In the attitude stabilization in turbulence problem, which is single loop on each axis, it was found that the error rate terms were not needed; in the air-to-air target tracking problem, it was found that while error rate terms were needed on both axes, the rate coefficient required on the multiloop axis was two orders of magnitude larger than that required on the single loop axis. These results raise two fundamental questions:

1. Are urgency function error rate terms ever required for single-loop tasks?
2. If so, how do they improve pilot model performance?

In order to answer these questions, a two-axis compensatory tracking task was employed. The command tracking signals were generated from uniformly distributed white noise as shown in Figure 1.

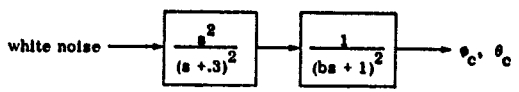


Figure 1. Command Tracking Signal Generation

Two sets of dynamics and two command tracking filter bandwidths were combined to produce six symmetric and asymmetric two-axis configurations as shown in Figure 2.

A flow diagram of the total piloted system appears in Figure 3. The pilot compensations were modeled using a gain K, lead T_L, and delay τ on each axis and programmed as follows:

$$\delta a = \text{Delay } (\tau)\left\{K_\phi(\phi_e + T_{L_\phi}\dot{\phi}_e)\right\}$$

$$\delta e = \text{Delay } (\tau)\left\{K_\theta(\theta_e + T_{L_\theta}\dot{\theta}_e)\right\}$$

The delays were fixed at $\tau = 0.3$ second, and the gains and leads were chosen to produce optimum model tracking statistics.

Configuration	Lateral Dynamics	Longitudinal Dynamics	b
1	$\frac{1}{s^2}$	$\frac{1}{s^2}$	8
2	$\frac{1}{s^2}$	$\frac{1}{s^2}$	4
3	$\frac{1}{s^2}$	$\frac{1}{s(s-1)}$	8
4	$\frac{1}{s^2}$	$\frac{1}{s(s-1)}$	4
5	$\frac{1}{s(s-1)}$	$\frac{1}{s(s-1)}$	8
6	$\frac{1}{s(s-1)}$	$\frac{1}{s(s-1)}$	4

Figure 2. Configurations Simulated for Command Tracking Task

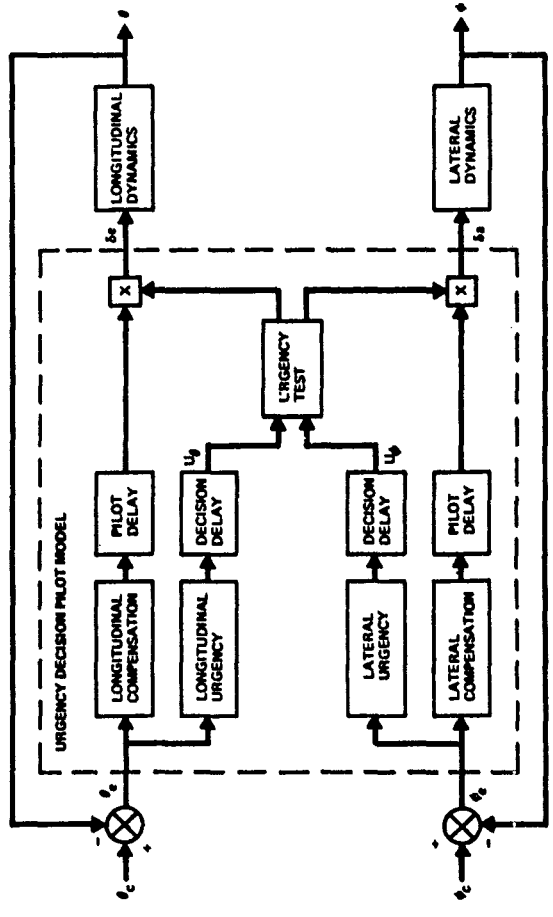


Figure 3. Control Configuration for Command Tracking Task

The urgency functions for this problem are of the form:

$$U_\phi = \left| \alpha_\phi \left| \phi_e \right| + \beta_\phi \frac{\phi_e}{\left| \phi_e \right|} \dot{\phi}_e \right|$$

$$U_\theta = \left| \alpha_\theta \left| \theta_e \right| + \beta_\theta \frac{\theta_e}{\left| \theta_e \right|} \dot{\theta}_e \right|$$

Since the planned simulation called for identical display gains on each axis, the relative position urgencies as perceived by the pilot are equal, so that α_ϕ and α_θ can be set to unity. The reflection of objective urgency as presented to the pilot by the display, in the subjective urgency adopted by the pilot is addressed in Reference 2. For the analysis of the complete model, the rate coefficients β_ϕ and β_θ are optimized along with the compensation gains and leads.

In order to discuss how the model is optimized for this problem, it is first necessary to describe the display used in the flight simulation. The CRT display consisted of a bright dot against a darkened background containing white crosshairs. Vertical displacement of the dot away from the origin represented θ_c, while horizontal displacement indicated ϕ_c. The pilot was instructed to keep the dot as close to the center as possible. This radial tracking error

$$r(t) = \sqrt{\phi_c^2(t) + \theta_e^2(t)}$$

is shown in Figure 4.

Explicitly, the quantity that the pilot and the pilot model optimized is given by

$$\bar{r} = \frac{1}{T} \int_0^T r(t)\, dt$$

In order to gain insight into the role played by error rate terms in the urgency functions, the model was optimized using a gradient method to select K_ϕ, K_θ, T_{L_ϕ}, T_{L_θ}, β_ϕ, and β_θ. The resulting tracking performance then could be compared with the model performance obtained by setting the β_ϕ and β_θ urgency rate coefficients to zero and optimizing K_ϕ, K_θ, T_{L_ϕ}, and T_{L_θ}.

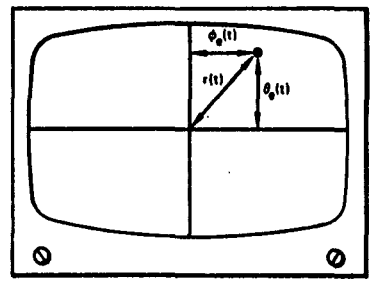


Figure 4. Flight Simulation Display

Since it was planned that the pilot in the subsequent fixed base flight simulation would fly a series of twenty thirty-second runs for a total of 600 seconds of data for each configuration, the model was optimized with respect to the average radial error for the same test schedule. In order to reduce statistical fluctuations of the command tracking signal, the command tracking signal sequence for each of the short runs was computed and scaled beforehand to produce a zero mean and unit standard deviation, i.e., unit command intensity.

The optimum model tracking scores and the corresponding model parameters for each of the six configurations with and without error rate terms in the urgency functions are given in Figure 5. Data on dwell fractions (the percentages of time the model controls each axis) and mean dwell times (the average length of individual control episodes) are presented in Figure 6.

Figure 5 shows that the model without rate terms cannot approach the tracking performance of the model with rate terms. In general the optimized incomplete model employs lower gains and higher leads than the optimized complete model. Figure 6

Configuration	Model with Error Rate Terms								Model without Error Rate Terms							
	$\bar{r}$	σ_r	K_ϕ	T_{L_ϕ}	β_ϕ	K_θ	T_{L_θ}	β_θ	$\bar{r}$	σ_r	K_ϕ	T_{L_ϕ}	β_ϕ	K_θ	T_{L_θ}	β_θ
1	0.299	0.0744	2.0	2.6	1.0	2.0	2.6	1.0	0.554	0.171	1.6	2.4	0	1.6	2.4	0
2	0.374	0.0866	1.8	2.6	1.0	1.8	2.6	1.0	0.726	0.211	1.4	2.4	0	1.4	2.4	0
3	0.464	0.122	2.2	2.0	1.0	2.0	2.0	1.0	1.20	0.452	1.4	2.4	0	1.6	2.2	0
4	0.578	0.138	2.2	2.0	1.0	2.0	2.0	1.0	1.60	0.491	1.2	2.6	0	1.6	2.4	0
5	0.647	0.178	2.2	1.8	1.0	2.2	1.8	1.0	5.81	3.29	0.4	8.0	0	0.4	8.0	0
6	0.725	0.158	2.6	2.0	4.0	2.6	2.0	4.0	6.80	2.79	0.4	8.0	0	0.4	8.0	0

Figure 5. Pilot Model Data for Command Tracking Task

Configuration	Model with Error Rate Terms				Model without Error Rate Terms			
	Dwell Fractions		Mean Dwell Times (seconds)		Dwell Fractions		Mean Dwell Times (seconds)	
	lat.	long.	lat.	long.	lat.	long.	lat.	long.
1	0.509	0.491	0.392	0.372	0.501	0.499	0.770	0.748
2	0.496	0.504	0.383	0.383	0.486	0.514	0.812	0.836
3	0.422	0.578	0.371	0.503	0.317	0.683	0.505	1.06
4	0.414	0.586	0.362	0.505	0.316	0.684	0.475	0.998
5	0.490	0.510	0.439	0.451	0.522	0.478	1.46	1.26
6	0.495	0.505	0.388	0.394	0.479	0.521	1.31	1.38

Figure 6. Dwell Fraction and Mean Dwell Time Data

shows that the incomplete model spends on the average much more time on each control episode than does the complete model, and on the asymmetric tasks spends a greater portion of its time controlling the unstable dynamics $1/s(s - 1)$.

A piloted fixed-base flight simulation was performed to verify the pilot model predictions. In order to limit the number of model parameters involved in the problem, display and controller effects in the flight simulation were minimized. This was done by using a large CRT display so that display motions were amplified enough to eliminate visual threshold effects and to present necessary rate information; controller effects were reduced by using a side-arm controller that had low force gradients and low but conspicuous breakout forces. The same computer programs, running on the same digital computers, were used for the simulation that were used for the model analysis work. For the flight simulation, the program operated in real-time and branched around the pilot model routine, reading the pilot's stick commands through analog-to-digital converters.

The pilot was instructed to keep the dot as close to the center of the screen as possible. This proved to be a difficult task, requiring several hours of training before the pilot achieved his optimum performance. It also required much experimentation with the controller and display scalings to arrive at a combination for each configuration that allowed the pilot to perform optimally.

The pilot flew thirty-second runs, in sets of ten or twenty. For each configuration, the best set of twenty (or the best pair of sets of ten) was saved, giving a total of 600 seconds of data. On the asymmetric configurations, the pilot flew ten runs with $1/s^2$ lateral and $1/s(s - 1)$ longitudinal, and ten with $1/s^2$ longitudinal and $1/s(s - 1)$ lateral.

The pilot's tracking scores for the six configurations, along with the scores for the complete model with urgency function error rate terms, are given in Figure 7. The excellent agreement between the tracking scores predicted by the complete model with the flight simulation tracking scores is shown in Figure 8.

A few comments concerning the large standard deviations σ_r in Figure 7 are in order. The precomputing and scaling of each command tracking signal sequence removed one source of variability from the runs by fixing the mean of ϕ_c and θ_c at zero with a standard deviation of one. However, this did nothing to standardize the frequency content of the command tracking signals. Owing to the short length of each run, the spectral content of the command sequences varied widely from run to run.

Configuration	Pilot		Model	
	$\bar{r}$	σ_r	$\bar{r}$	σ_r
1	0.311	0.0250	0.299	0.0744
2	0.420	0.0619	0.374	0.0866
3	0.506	0.0778	0.464	0.122
4	0.609	0.0678	0.578	0.138
5	0.565	0.0882	0.647	0.178
6	0.650	0.0952	0.725	0.158

Figure 7. Command Tracking Flight Simulation and Model Data

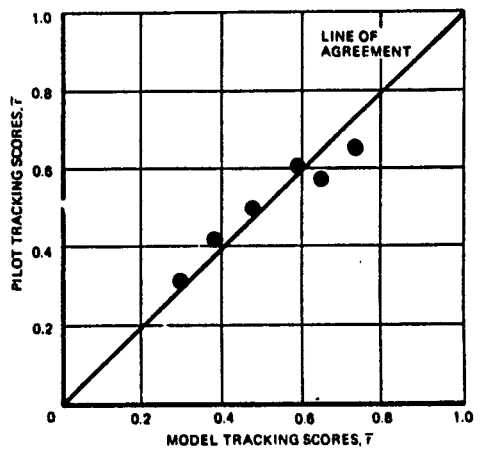


Figure 8. Agreement of Model and Flight Simulation Tracking Scores

This is illustrated by Figure 9, which presents time histories of two command tracking sequences from the simulation; each sequence represents thirty seconds of real time. It is likely that this contributed to the variability in both pilot and model performance.

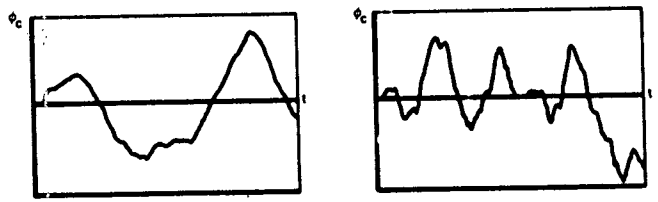


Figure 9. Sample Command Tracking Command Time Histories

It can also be seen in Figure 7 that the standard deviations for the model are generally larger than those for the pilot. It was sometimes the case that initial transient command tracking errors and rates were of such a nature that the pilot had to abort a run shortly after it began (e.g., the dot started near the edge of the screen and quickly went out of sight). Such runs were not counted, and no statistics were kept for them. However, the model did not have this luxury and had to fly every command sequence that came along. It is likely that a few of these runs with large initial transients in each set increased the standard deviations of the model's tracking scores.

CONCLUSIONS

The Northrop urgency decision pilot model was applied to a compensatory tracking task involving marginally stable and unstable second order system dynamics in an effort to answer two fundamental questions concerning the model's urgency function:

1. Are urgency function error rate terms ever required for single loop tasks?
2. If so, how do they improve pilot model performance?

Comparison of the flight and pilot model simulations demonstrated the complete model's ability to predict the pilot's tracking statistics, just as it demonstrated that the incomplete model was incapable of such performance. Analysis using the pilot model with and without the benefit of error rate information in the urgency functions demonstrated the importance of this error rate information in the following ways:

1. The complete model agreed well with the flight simulation, whereas the incomplete model had badly degraded performance not seen in the flight simulation.
2. The error rate information in the complete model led to attention shifting rates necessary to control the unstable systems, while the incomplete model was not able to initiate corrective action promptly enough to maintain low error rates.
3. Since large rates built up, the incomplete model was forced to adopt higher leads to control them, while adopting lower gains to avoid overcontrolling the system.
4. In asymmetric tasks, the incomplete model was forced to spend a disproportionate amount of time trying to control the rates generated by the less stable dynamics.

From the results of this study, it is reasonable to conclude that urgency function error rate terms are required for analysis of any unstable or marginally stable system, or in general, for any system capable of evolving appreciable error rates.

REFERENCES

1. Onstott, E. D., Multi-Axis Pilot-Vehicle Dynamics, Proceedings of the Tenth Annual Conference on Manual Control, AFIT/AFFDL, April 1974.
2. Onstott, E. D., Task Interference in Multi-Axis Aircraft Stabilization, Proceedings of the Twelfth Annual Conference on Manual Control, University of Illinois, May 1976.
3. Onstott, E. D., and Faulkner, W. H., Prediction of Pilot Reserve Attention Capacity During Air-to-Air Target Tracking, Proceedings of the Thirteenth Annual Conference on Manual Control, MIT, June 1977.

Session II

PERFORMANCE, ATTENTION ALLOCATION AND MENTAL LOAD

Chairman: W. H. Levison

Research Publication GMR-2394

THE APPLICATION OF INTEGRAL PERFORMANCE CRITERIA TO THE ANALYSIS OF DISCRETE MANEUVERS IN A DRIVING SIMULATOR

Brian S. Repa
Robert S. Zucker
Engineering Mechanics Department
General Motors Research Laboratories
Warren, Michigan 48090

Walter W. Wierwille
Consultant (Virginia Polytechnic
Institute and State University)
Blacksburg, Virginia

March 31, 1977

To be presented at the Thirteenth Annual Conference on Manual Control in Cambridge, Massachusetts, June 15, 1977.

PURPOSE

The purpose of this study was to investigate the influence of vehicle transient response characteristics on driver-vehicle performance in discrete maneuvers as measured by integral performance criteria.

SUMMARY

A group of eight ordinary drivers was presented with a series of eight vehicle transfer function configurations in the Virginia Polytechnic Institute and State University driving simulator. Performance in two discrete maneuvers was analyzed by means of integral performance criteria. The following results were obtained:

1. The step wind gust disturbance regulation task was found to be more challenging and more sensitive to differences in vehicle characteristics than a task requiring correction of an artificially induced lane position error. Subsequent remarks apply to results obtained with the step wind gust task.

2. In comparing the various integral criteria, it was observed that the mean scores for all subjects for the Integral of Squared Time Multiplied by Squared Error (ISTSE) criterion showed the greatest separation among vehicle configurations.

3. Steering wheel angle was found to be the most sensitive output when used with this criterion, with yaw angle and lane position being less effective in that order.

4. Performance as measured by the ISTSE criterion showed the following effects for the various vehicle yaw rate transfer function parameters:

   - A natural frequency of 1.75 rad/sec resulted in significantly degraded scores in comparison with frequencies of 3.5, 7.0, and 14.0 rad/sec. The higher the natural frequency, the quicker the vehicle response.
   - A damping ratio of 1.6 resulted in a significantly higher integral score for steering wheel angle than did damping ratios of 0.4 and 0.8. Lateral position and yaw angle measures were insensitive to differences in this parameter. Higher damping ratios result in less vehicle overshoot in response to a step steering wheel input, but they also produce longer response times.

- Setting the yaw lead time constant to zero significantly degraded the integral performance scores when compared with time constants of 0.28 and 0.57 sec. Adjustments to the lead time constant directly affected the rate of onset of the vehicle yaw rate response, with larger time constants producing quicker responses.
- The compound vehicle response parameter of response time divided by "effective" damping ratio was found to be highly correlated with the integral measure of lane position error.

Future efforts will be directed toward determining if integral measures, such as ISTSE, relate to subjectively-sensed problems of discrete control in full-scale testing, in addition to being discriminatory across different vehicle characteristics.

INTRODUCTION

There is a substantial background in the driving research literature relating to discrete maneuvers, although much of the material has been qualitative, e.g., [1-5]. There is also a considerable body of research in the automatic control field on the selection of integral performance measures of transient responses which lead to "satisfactory-appearing" responses, e.g., [6-8]. To accomplish the guidance and control functions of driving, the driver sets up a variety of closed loops about the vehicle, each of which depends on a variable such as lateral position or yaw angle [9]. By analogy, a satisfactory driver-vehicle system should mimic certain of the dynamic features of "good" automatic control systems. The goal here, then, is to find integral measures of one or several driver-vehicle system outputs which will lead to "satisfactory" handling responses. The motivation is that a single index of performance could be used instead of a recorded time history.

This report presents the effects of changes in vehicle control dynamics on driver-vehicle performance in discrete tasks as measured by integral performance criteria. The independent variables were the vehicle dynamics as defined by the yaw rate to steering wheel transfer function. Different configurations were obtained by changing one at a time either yaw rate natural frequency, damping ratio, or lead time constant. Two transient maneuvers were investigated; namely, step wind gust disturbance regulation and compensation for an artificially induced lane position error. The study was performed in a moving base driving simulator. Transient steering wheel inputs and vehicle motions were analyzed from strip chart recordings, and on this basis a variety of integral performance measures were evaluated.

This study is part of the overall efforts of the Driver-Vehicle Performance Project which is aimed at determining the influence of vehicle control characteristics on driving performance and comfort. The philosophy is to employ driving simulators as an experimental prelude to full-scale investigations, particularly in the development of the tests and measures required to meet the project's objectives.

DRIVING SIMULATOR FACILITY

The investigation was conducted on the Virginia Polytechnic Institute and State University (VPI & SU) driving simulator. This facility features a three degree of freedom motion platform (yaw, roll, and lateral translation) and a computer generated roadway scene. Closed-loop control of the simulated vehicle by the human subject is accomplished by sending signals from steering wheel, accelerator, and brake pedal transducers to the simulated equations of motion which, in turn, modify the platform motion and roadway scene. More complete descriptions of the simulator are found in [10-12].

The method used for representing different vehicles on the simulator involves the derivation of transfer functions that closely match the time domain responses measured for actual vehicles. The contributions of vehicle, tire, and suspension characteristics are accounted for in terms of transfer function parameters. The transfer function that closely approximates a vehicle's yaw rate response is given by [13]:

$$\frac{r}{\delta_{sw}}(s) = \frac{T_r s + 1}{\frac{1}{\omega_n^2} s^2 + \frac{2\zeta}{\omega_n} s + 1} \times \left(\frac{r}{\delta_{sw}}\right)_{ss}$$

where r = yaw rate, deg/sec

δ_{sw} = steering wheel angle, deg

T_r = lead time constant, sec

ω_n = natural frequency, rad/sec

ζ = damping ratio

$\left(\frac{r}{\delta_{sw}}\right)_{ss}$ = steady state yaw rate gain, $\frac{\text{deg/sec}}{\text{deg}}$.

Eight experimental yaw rate transfer functions were chosen for investigation. Table 1 lists the specific parameter values employed for the different configurations as well as the values for other characteristics that were held fixed. The different configurations were partitioned into three comparison groups to determine the individual effects of changes in natural frequency, damping ratio, and lead time constant.

PROCEDURE

A group of eight subjects, ranging in age from 21 to 32 years and drawn from students and personnel at Virginia Polytechnic Institute

and State University, were tested on each of the eight vehicle configurations. Two of the subjects were female. Each subject held a valid driver's license and had a minimum of three years of driving experience. None of the subjects had previously performed in the driving simulator. The order in which the configurations were presented was randomized by a Latin Square Design.

Upon arrival, each subject was asked to read the instructions (Appendix A) and was then seated in the simulator for further verbal instructions concerning its operation. Figure 1 shows a time line diagram of the events of a particular run. Following a one-minute practice period and prior to the discrete maneuvers, subjects performed in a random wind gust disturbance regulation task which was treated in a previous report [12].

DISCRETE MANEUVERS

Two discrete tasks were employed: compensation for an artificially induced lane position error and regulation of a step wind gust disturbance. One of the advantages of using a simulator is the ability to introduce artificial changes in vehicle output variables. Figure 2 shows time traces for one subject for different vehicle conditions following step changes in lateral position. The discontinuity on lateral position, 90 cm (3 ft) is readily apparent in the time trace and served as an unexpected error to be corrected by the subjects. Time traces on one subject for the step wind gust regulation task are shown in Fig. 3. This task differs from the previous one in that motion cues are present along with a steady side force.

PERFORMANCE MEASURES

The time traces shown in Figs. 2 and 3 appeared to be quite amenable to integral transient response measures that would provide a single index of goodness representative of an entire time history. The underlying difficulty was the choice of the best measure to use for these particular cases. One performance index will emphasize some properties of the response more than others. In addition, the performance index may be chosen so that only one or several system outputs affect its value. In the driver-vehicle system more than one output must be examined. Path and heading errors as well as steering wheel inputs should be considered, the former as measures of performance and the latter as a measure of driver effort or cost.

There have been various proposals and suggestions of specific integral measures in the literature, e.g., [6-8]. The ones that were examined include:

(1) Integral of absolute error (IAE)

$$I_1 = \int_0^T |e|dt$$

(2) Integral of squared error (ISE)

$$I_2 = \int_0^T e^2 dt$$

(3) Integral of time multiplied by absolute error (ITAE)

$$I_3 = \int_0^T t|e|dt$$

(4) Integral of time multiplied by squared error (ITSE)

$$I_4 = \int_0^T t\, e^2 dt$$

(5) Integral of squared time multiplied by squared error (ISTSE)

$$I_5 = \int_0^T t^2 e^2 dt$$

(6) Integral of squared time multiplied by absolute error (ISTSE)

$$I_6 = \int_0^T t^2 |e|dt$$

A geometric interpretation of the ITAE criterion is shown in Fig. 4.

Three additional and somewhat unconventional measures involving multiplication by inverse time were also included for the purpose of giving additional emphasis to the initial aspects of the transient responses.

(7) Integral of inverse time multiplied by absolute error (IITAE)

$$I_7 = \int_{a>0}^T t^{-1} |e|dt$$

(8) Integral of inverse time multiplied by squared error (IITSE)

$$I_8 = \int_{a>0}^T t^{-1} e^2 dt$$

(9) Integral of inverse squared time multiplied by absolute error (IISTAE)

$$I_9 = \int_{a>0}^{T} t^{-2} \, |e|dt$$

Numerical values for the various performance measures were obtained by digitizing time traces of lateral position, yaw angle, and steering wheel angle on a Bendix Graphscan Digitizer and then performing the required integrations digitally. An integration time of 10 seconds was used for the lateral position jump task and 14 seconds for the step wind gust task.

As a means of comparison with the various integral measures, peak lane position overshoot for the step wind gust task was also selected as a measure of performance, primarily because of its intuitive safety relevance.

DRIVER-VEHICLE PERFORMANCE RESULTS

Step Wind Gust Disturbance Regulation Task

Table 2 contains a summary of the statistical test results for the peak lane position overshoot measure. Application of Duncan's New Multiple Range Test [17] showed that in comparisons of natural frequency levels, Configuration #4 (ω_n = 1.75 rad/sec) was significantly different from the other three configurations (ω_n = 3.5, 7.0, and 14.0 rad/sec), and when comparing variations in lead time constant, that Configuration #8 (T_r = 0) was significantly different from Configurations #1 (T_r = 0.28 sec) and #7 (T_r = 0.57 sec). No significant differences were observed as a result of variations in damping ratio.

To make a comparison between the various integral criteria under consideration, the criterion scores are normalized and plotted versus natural frequency for lateral position, yaw angle, and steering wheel angle transients in Figs. 5-7. Also included for comparison in Fig. 5 for measures of lateral position error are Peak Lateral Position Overshoot Squared. To avoid unnecessary cluttor, integral measures I_6-I_9 are omitted from the figures. These measures showed less discrimination across natural frequency than either of the top three criteria, ISE, ITSE, and ISTSE.

Because of its greater spread in scores across configurations, ISTSE was selected for closer examination. Table 3 summarizes the statistical test results for this measure, and Figs. 8-10 show the influences of variations in natural frequency, damping ratio, and lead time constant as determined by this measure. The mean scores for the eight subjects with the corresponding standard deviations are indicated.

Integral measures of yaw angle and steering wheel angle show the most discrimination across configurations. Application of Duncan's New Multiple Range Test to the data revealed the following:

> Natural frequency comparison: Performance for the configuration with a natural frequency of 1.75 rad/sec was significantly degraded from the others.
>
> Damping ratio comparison: The high damping configuration (ζ = 1.6) produced a significantly higher integral measure of steering wheel angle than the others.
>
> Lead time constant comparison: Performance for the configuration with T_r = 0 was significantly less than for the other configurations.

In an attempt to compare performance on all of the configurations in terms of a common and relevant response characteristic, the compound parameter of response time divided by "effective" damping ratio was employed. To do this, the damping ratio of the second order system without lead (i.e. having no transfer function "zero") that would give the same value of yaw rate overshoot that was observed for each of the configurations was assigned as the "effective" damping ratio. As shown in Fig. 11, a remarkably close fit to a straight line results when the Integral of Squared Time Multiplied by Squared Lateral Position Error is plotted as a function of this compound parameter. The Pearson Product Moment Correlation Coefficient corresponding to this relationship is 0.961 which is highly significant ($p < 0.001$).

COMPENSATION FOR ARTIFICIALLY INDUCED LANE POSITION ERROR

A comparison between the various integral criteria using steering wheel angle as the error signal is shown in Fig. 12 for the four natural frequency configurations. Because subjects varied widely in the time they took to initiate a correction for the step change in lane position it was decided that an explicit time penalty in the cost functional would be inappropriate. In addition, measures with explicit time, i.e., ITSE, ITAE, etc., resulted in a poorer performance score for the vehicle with a natural frequency of 7.0 rad/sec than for the one with 3.5 rad/sec which did not appear to reflect an accurate ranking of the vehicles. For these reasons, the ISE measure was chosen for a more complete analysis of the data.

Table 4 summarizes the statistical test results for this measure, and Fig. 13 illustrates the influence of variations in the three vehicle response characteristics as determined by the measure for steering wheel angle. Measures of lateral position error were insensitive to differences in vehicle configuration. ISE measures of yaw angle and steering wheel angle did show statistically significant differences between different levels of damping ratio and lead time constant but not for natural frequency, although the latter just missed being significant at the $p = 0.10$ level. Duncan's New Multiple Range Test again showed that a value of

ζ = 1.6 was significantly different from the other levels of damping ratio and that T_r = 0 was significantly different from the other two levels of lead.

DISCUSSION OF RESULTS

Of the two discrete maneuvers employed in the present study, the step wind gust disturbance task was the more challenging and discriminating among vehicle configurations. Both visual and motion cues accompanied the disturbance, and unless rapid compensatory control on the part of the driver was taken, the vehicle would have been forced off the road. Introduction of a simple step change in lane position, however, resulted in widely varying responses on the part of the subjects. Subjects were not given specific instructions on this task other than the general instruction to make whatever corrections were necessary so as to maintain their normal highway position. Since the artificially induced lane position error was only 90 cm (3 ft) in magnitude, it may not have been sufficient to motivate an immediate correction on the part of the subjects. It is possible that more specific instructions or a larger induced error would make the task more sensitive to differences in vehicle configurations.

Integral measures of the various driver-vehicle system outputs appear to offer an effective means for comparing transient performance with different vehicle response configurations. The fact that the ISTSE measure was shown to discriminate among the different configurations on a statistical basis using only a single trial for each driver-vehicle combination is noteworthy. Comparison of the mean scores of the various performance measures across vehicle configurations must be done with some caution, however, since it does not take into account the variability inherent in a given measure. Thus, while Fig. 5 shows that mean scores for ISTSE appear much more discriminating than those for Peak Lane Position Overshoot (PLPO), for example, statistical tests with the two measures, Tables 2 and 3, indicate similar discrimination capabilities. Intuitively, the integral measure should be superior because it is determined by the complete transient response rather than a specific aspect of it. The peak overshoot appears to be the predominant feature in the lateral responses in this instance, however.

Findings on the effects of changes in the vehicle response characteristics are in good agreement with previously reported qualitative results. This result provides added support for the integral criteria. Linke, Richter, and Schmidt [1] found that the higher the resonant frequency of the vehicle, the better the mean subjective rating as determined from a group of unskilled drivers performing a high speed lane change maneuver. Although these authors discounted the importance of yaw damping, inspection of two of their vehicles with similar resonant frequencies reveals that the one with the least damping and hence, lowest response time was judged the most favorably. In its report to the International Standards Organization [2], Sweden showed a correlation coefficient of 0.926 between subjective ratings in a lane change maneuver and time lags between steering wheel inputs and yaw velocity responses. In terms of the parameters used in our study, the initial slope of the yaw rate response is given by

$$\dot{r}(o) = \left(\frac{r}{\delta_{sw}}\right)_{ss} \times T_r \times \omega_n^2 \times \delta_{sw} \quad .$$

In a real vehicle, factors that cause an increase in T_r generally bring about a decrease in ω_n and as a rough approximation, $T_r \approx 1/\omega_n$. The time lag between steering wheel angle and yaw rate, as used by the ISO, can be roughly approximated by steady state yaw rate divided by the slope. Thus,

$$\text{Time lag} \approx \frac{r_{ss}}{\dot{r}(o)} = \frac{r_{ss}}{\left(\frac{r}{\delta_{sw}}\right)_{ss} \times \omega_n \times \delta_{sw}} = \frac{1}{\omega_n} \quad .$$

If the ISO data are replotted in terms of natural frequency, the curve shown in Fig. 14 results which bears a striking resemblance to our own findings.

For transient maneuvers including double lane changes, unexpected obstacle avoidance, and step disturbance regulation, STI [5] has recently reported that natural frequencies below 3 rad/sec may represent a lower boundary for acceptable performance and that based on driver opinion ratings, natural frequency appears more important than damping ratio.

The effects of the vehicle characteristics on step wind gust performance are also very similar to those described in a previous GMR report on random wind gust disturbance regulation [12]. The significance levels are somewhat less for the present findings, but this is not at all surprising in view of the methodological differences. The present integral performance criteria were applied to the subjects' first exposure to a step gust disturbance with each vehicle, and the period of integration was on the order of 10 to 15 seconds. In the random gust disturbance case, subjects were allowed a one-minute practice period under gusting conditions before the 2½ minute data run was taken. The fact that the trends for the step gust disturbance case using integral performance criteria are as strong as indicated under the circumstances is justification for further use of this approach.

The work in the present study is based entirely on objective measures of directional control performance in a driving simulator, and the natural extension would be to correlate these measures with handling qualities ratings during discrete maneuvers under full-scale conditions. If the integral measures are found to relate to the subjectively sensed problems of discrete control in addition to being discriminatory among different vehicle conditions, then a reasonably simple means for measuring objective handling properties will be available. Such measures should lend themselves to easy data gathering and on-board mechanization. Steering wheel inputs are readily available, and approximate path and heading outputs can probably

be computed on-board via "washed-out" integrations of lateral acceleration and yaw rate. The resulting signals can then be squared, time weighted, and integrated up to 10 or 15 seconds by an on-board analog circuit which is reset at the start of each maneuver. The resulting objective scores, i.e., ISTSE, and corresponding opinion ratings can then be logged for later comparison. Since the pilot simulator data show that integral measures of yaw angle and steering wheel angle are more sensitive to changes in vehicle response characteristics than measures of path error, possible difficulties in taking the second integral of lateral acceleration to obtain lateral position error should not seriously compromise the analysis.

Previous attempts to objectively measure driver-vehicle performance in discrete maneuvers have typically used for a score the number of cones displaced from a course or the maximum speed through the course without hitting any cones. The presence of the cone course itself allows anticipation on the part of the driver and requires considerable time to set up and maintain. The use of an integral performance criterion does not have these difficulties. The GM Variable Response Vehicle [14] has the capability to simulate a step disturbance input so the step wind gust disturbance regulation task can be largely duplicated under full-scale conditions. The step jump in lane position is obviously a simulator novelty. However, under full-scale conditions, at some unexpected time set by the experimenter and in a direction chosen at random, the subject can be commanded to switch from a center lane to either the right or left lane as rapidly as possible. The integral criteria should be equally suitable for this type of maneuver. The extension of this procedure to full-scale testing should be relatively straightforward, and combined with subjective opinion data it will hopefully lead to an improved method for distinguishing "good" and "bad" handling vehicles.

REFERENCES

1. W. Lincke, B. Richter, and R. Schmidt, "Simulation and Measurement of Driver Vehicle Handling Performance," SAE Paper No. 730489, May 1973.

2. "Experimental Study of a Lane Change Manoeuvre with Mechanized Sinusoidal Steering Input," ISO/TC 22/SC 9 (Sweden-5) 52, October 1974.

3. W. Bergman, "Measurement and Subjective Evaluation of Vehicle Handling," SAE Paper No. 730492, May 1973.

4. W. Bergman, "Considerations in Determining Vehicle Handling Requirements," SAE Paper No. 690234, January 1969.

5. D. T. McRuer and R. H. Klein, Automobile Controllability -- Driver/Vehicle Response for Steering Control, Volume 1, Summary Report, Systems Technology, Inc., Contract No. DOT-HS-359-3-762, February 1975.

6. D. Graham and R. C. Lathrop, "The Synthesis of Optimum Response: Criteria and Standard Forms," Trans. AIEE, Vol. 72, Part II, November 195.. 273-288.

7. W. C. Schultz and V. C. Rideout, "The Selection and Use of Servo Performance Criteria," Trans. AIEE, Vol. 76, Part II, January 1958, 383-388.

8. J. E. Gibson, et. al., "A Set of Standard Specifications for Linear Automatic Control Systems," Trans. AIEE, Vol. 80, Part II, May 1961, 65-74.

9. D. H. Weir and D. T. McRuer, "Models for Steering Control of Motor Vehicles," Fourth Annual NASA-University Conference on Manual Control, The University of Michigan, March 1968.

10. W. W. Wierwille, "A Part-Task Driving Simulator for Teaching and Research," Computers in Education for ASEE Transactions, December 1973.

11. W. W. Wierwille, "Driving Simulator Design for Realistic Handling," Proceedings of the Third International Conference on Vehicle System Dynamics, Swets and Zeitlinger B.V., Amsterdam, 1975.

12. B. S. Repa and W. W. Wierwille, "Driver Performance in Controlling a Driving Simulator with Varying Yaw Rate Characteristics," GM Research Publication GMR-2205, July, 1976.

13. R. T. Bundorf and R. L. Leffert, "The Cornering Compliance Concept for Description of Directional Control (Handling) Properties", Engineering Publication 2771, GM Proving Ground.

14. K. J. McKenna, "A Variable Response Vehicle -- Description and Applications", GM Proving Ground Engineering Publication 5665, June, 1974.

15. H. O. Hartley, "The Maximum F-Ratio as a Short Cut Test for Homogeneity of Variance", Biometrika, 1950, Vol. 37, 308-312.

16. G. Keppel, Design and Analysis, a Researcher's Handbook, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1973, 465-466.

17. A. L. Edwards, Experimental Design in Psychological Research, Rinehart and Company, Inc., New York, 1960, 136.

TABLE 1 - YAW RATE CONTROL RESPONSE PARAMETERS

Condition	$1/T_r$ (sec^{-1})	ζ —	ω_n (rad/sec)
1	3.5	0.8	3.5
2	14.0	0.8	14.0
3	7.0	0.8	7.0
4	1.75	0.8	1.75
5	3.5	0.4	3.5
6	3.5	1.6	3.5
7	1.75	0.8	3.5
8	—	0.8	3.5

Yaw Rate to Steering Wheel Input Gain: 0.21 $\frac{deg/sec}{deg}$

Control Sensitivity (Lateral Acceleration to Steering Wheel Input Gain): 0.92 $\frac{G}{100\ deg}$

Roll Compliance: 6.2 deg/G

Roll Natural Frequency: 6 rad/sec

Roll Damping Ratio : 0.50

TABLE 2 - ANALYSIS OF VARIANCE SUMMARY FOR PEAK LANE POSITION OVERSHOOT IN THE STEP WIND GUST REGULATION TASK

Source of Variation	Degrees of Freedom		Maximum Lane Deviation		
	Uncorrected	Corrected†	Sum of Squares	Mean Square	F-Ratio
Natural Frequency (ω_n)	3	-	18.8	6.26	2.91*
Error	21	-	45.2	2.15	
Damping Ratio (ζ)	2	-	2.10	1.05	0.506
Error	14	-	29.1	2.07	
Lead Time Constant (T_r)	2	1	105.8	52.9	13.76***
Error	14	7	53.8	3.84	

Significance Level: * = 0.1, ** = 0.05, *** = 0.01

† The Greenhouse-Geisser correction [16] was applied to the degrees of freedom for those variables not meeting the homogeneity of variance assumption using Hartley's test [15].

TABLE 3 - ANALYSIS OF VARIANCE SUMMARY FOR ISTSE CRITERION FOR THE STEP WIND GUST REGULATION TASK

Source of Variation	Degrees of Freedom		Lateral Position Error			Yaw Angle Error			Steering Wheel Angle		
	Uncorrected	Corrected†	Sum of Squares	Mean Square	F-Ratio	Sum of Squares	Mean Square	F-Ratio	Sum of Squares	Mean Square	F-Ratio
Natural Frequency (ω_n)	3	1	1.04×10^8	0.347×10^8	4.7*	6.02×10^5	2.01×10^5	10.53**	8.53×10^{10}	2.84×10^{10}	15.9***
Error	21	7	1.54×10^8	0.074×10^8		4.01×10^5	0.191×10^5		3.75×10^{10}	0.178×10^{10}	
Damping Ratio (ξ)	2	1	8.79×10^6	4.40×10^6	0.566	2.37×10^5	1.18×10^5	1.56	5.45×10^{10}	2.72×10^{10}	4.78*
Error	14	7	10.9×10^7	0.78×10^7		10.6×10^5	0.76×10^5		7.97×10^{10}	0.569×10^{10}	
Lead Time Constant (T_2)	2	1	7.01×10^8	3.5×10^8	6.93**	8.82×10^6	4.41×10^6	9.64**	4.14×10^{11}	2.07×10^{11}	14.45***
Error	14	7	7.07×10^8	0.51×10^8		6.40×10^6	0.457×10^6		2.00×10^{11}	0.143×10^{11}	

Significance Level: * = 0.10, ** = 0.05, *** = 0.01

† The Greenhouse-Geisser correction [16] was applied to the degrees of freedom for those variables not meeting the homogeneity of variance assumption using Hartley's test [15].

TABLE 4 - ANALYSIS OF VARIANCE SUMMARY FOR THE INTEGRAL OF SQUARED ERROR PERFORMANCE CRITERION FOR THE ARTIFICIALLY INDUCED LANE POSITION ERROR TASK

Source of Variation	Degrees of Freedom		Lateral Position			Yaw Angle			Steering Wheel Angle		
	Uncorrected	Corrected†	Sum of Squares	Mean Square	F-Ratio	Sum of Squares	Mean Square	F-Ratio	Sum of Squares	Mean Square	F-Ratio
Natural Frequency (ω_n)	3	1	663.6	221.2	0.841	14.8	4.94	0.940	1.30×10^6	4.35×10^5	3.50
Error	21	7	5525.0	263.1		110.5	5.26		2.61×10^6	1.24×10^5	
Damping Ratio (ξ)	2	1	633.5	316.7	1.618	65.7	32.9	4.67*	6.92×10^6	3.46×10^6	7.16**
Error	14	7	2740.0	195.8		98.5	7.04		6.76×10^6	4.83×10^5	
Lead Time Constant (T_2)	2	1	2140.0	1070.0	3.46	43.0	21.5	7.60**	1.74×10^6	8.69×10^5	15.61***
Error	14	7	4333.0	309.0		39.6	2.83		7.80×10^5	5.57×10^4	

Significance Level: * = 0.10, ** = 0.05, *** = 0.01

† The Greenhouse-Geisser correction [16] was applied to the degrees of freedom for those variables not meeting the homogeneity of variance assumption using

APPENDIX A

INSTRUCTIONS

The purpose of this experiment is to obtain a better understanding of normal driving behavior.

You will be seated in the driver's seat of an automotive mock-up. You will be presented with a visual display consisting of a moving, geometrical roadway simulation and a dashboard speedometer. During operation of the simulator, you will experience simulated vehicle motions corresponding to the driving conditions and your control maneuvers. Your control of the simulator's speed and road position will be by means of a standard steering wheel and accelerator pedal as in a normal automobile. After being seated on the platform you will be given instructions by, and may communicate with the experimenter via the dash mounted (upper right) speaker/microphone.

The total experiment will take approximately 1 hour and 15 minutes to complete. There will be 8 experimental runs of approximately 7 minutes each. There will be a brief intermission between each experimental run. After 4 runs, you will be allowed to leave the simulator for a 5 minute rest. Thereafter, you will return to the simulator and the last four runs will be completed.

During the run you are to perform two tasks:

1. maintain a speed of 55 mph, and
2. maintain a normal right-lane highway driving position.

Please keep in mind that you are at all times to drive as you normally would on a highway.

During the course of the experiment, inputs will be artifically introduced into the simulation, causing the vehicle to deviate in different ways. Your job will be to make corrections for these deviations so as to maintain the normal highway position previously established. (Note: At the beginning of the experiment you are to accelerate to 55 mph.) Shortly thereafter deviations in the simulation will occur.

The experimental precedure will be as follows:

1. Be seated in the driver's seat; adjust seat position and fasten safety belt.
2. Become familiar with controls, speaker/microphones, and emergency motion cut-off button.

   NOTE: Activation (1 push) of the emergency motion cut-off button halts all motion of the simulator platform. If at any time during the experiment you sincerely fee that continued simulator operation would not be agreeable with you, please verbally notify the experimenter and depress (once), the emergency motion cut-off button. You may leave the platform (to the left only) if and only if all platform motion has stopped.

3. Communications checkout and questions.

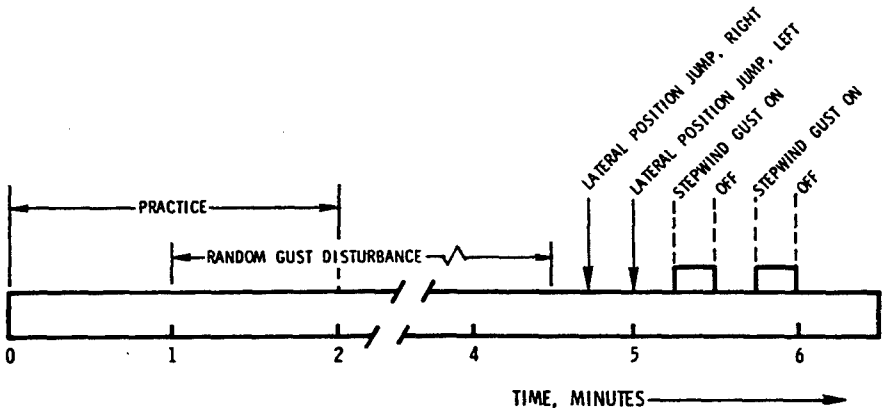


FIG. 1 SCHEDULE OF EVENTS

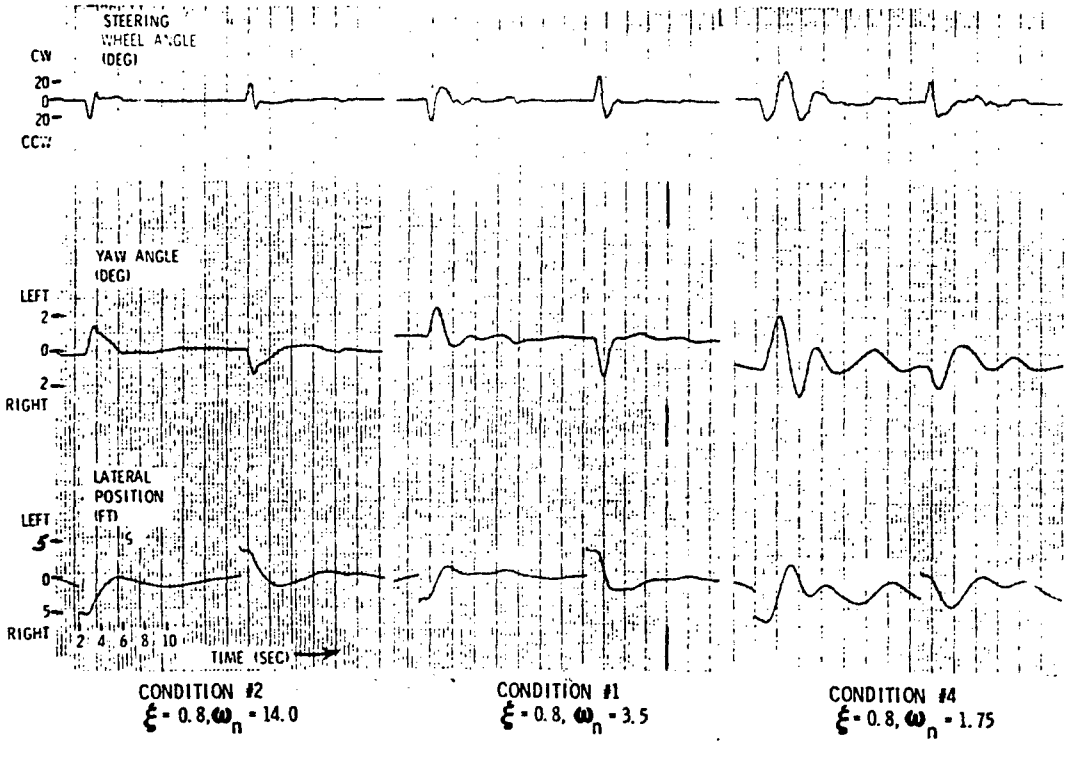


FIG. 2 TIME TRACES FOR THE ARTIFICIALLY INDUCED LANE POSITION ERROR TASK WITH DIFFERENT LEVELS OF NATURAL FREQUENCY

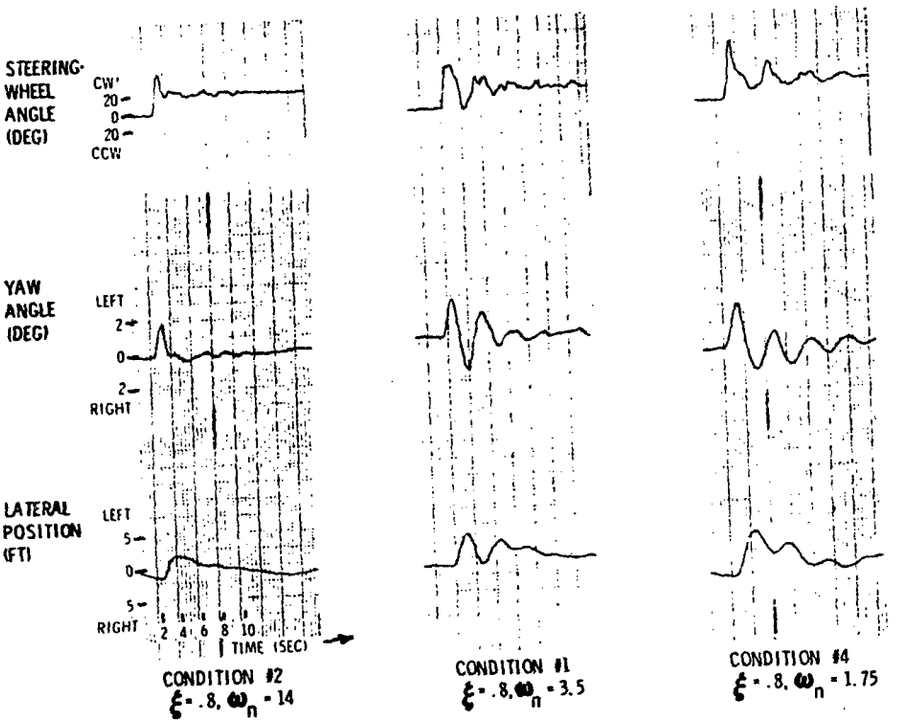


FIG. 3 TIME TRACES FOR STEP WIND GUST REGULATION FOR DIFFERENT LEVELS OF NATURAL FREQUENCY

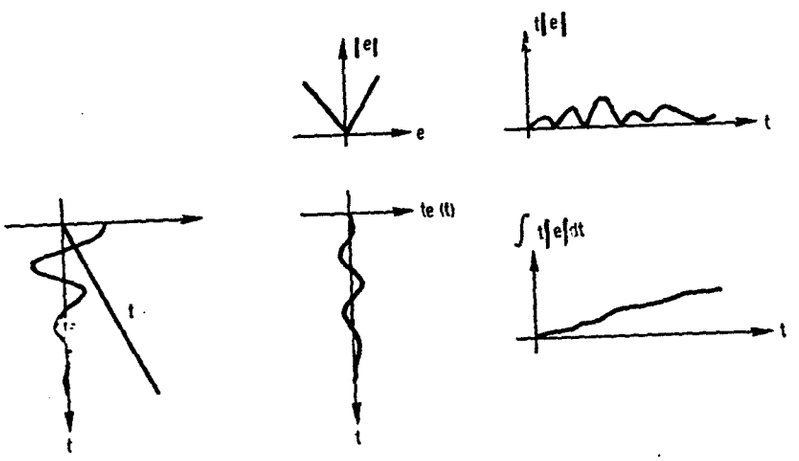


FIG. 4 GEOMETRIC INTERPRETATION OF INTEGRAL OF TIME MULTIPLIED BY ABSOLUTE ERROR (ITAE) CRITERION [7]

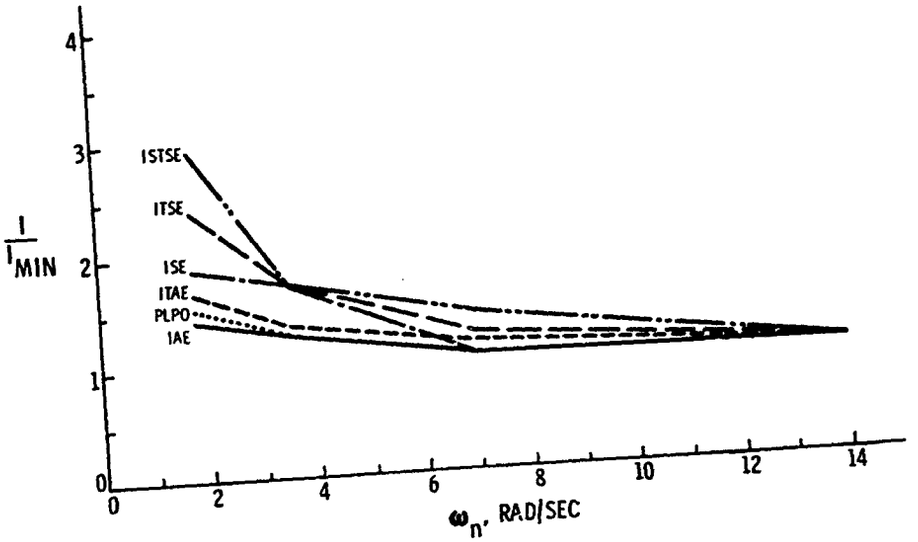


FIG. 5 COMPARISON OF PERFORMANCE CRITERIA AS A FUNCTION OF NATURAL FREQUENCY, WITH LATERAL POSITION AS THE ERROR

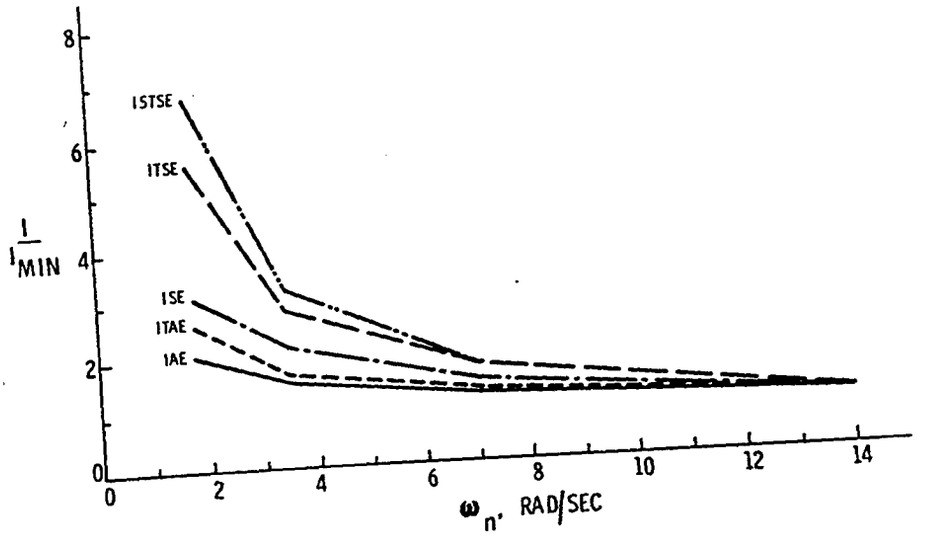


FIG. 6 COMPARISON OF PERFORMANCE CRITERIA AS A FUNCTION OF NATURAL FREQUENCY, WITH YAW ANGLE AS THE ERROR

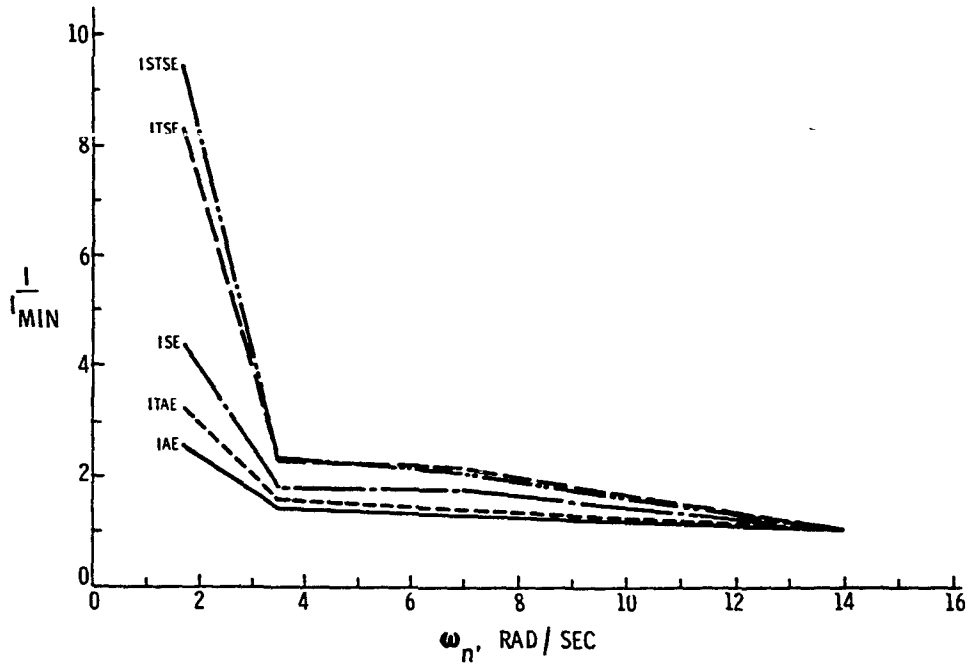


FIG. 7 COMPARISON OF PERFORMANCE CRITERIA AS A FUNCTION OF NATURAL FREQUENCY, WITH STEERING WHEEL ANGLE AS THE ERROR

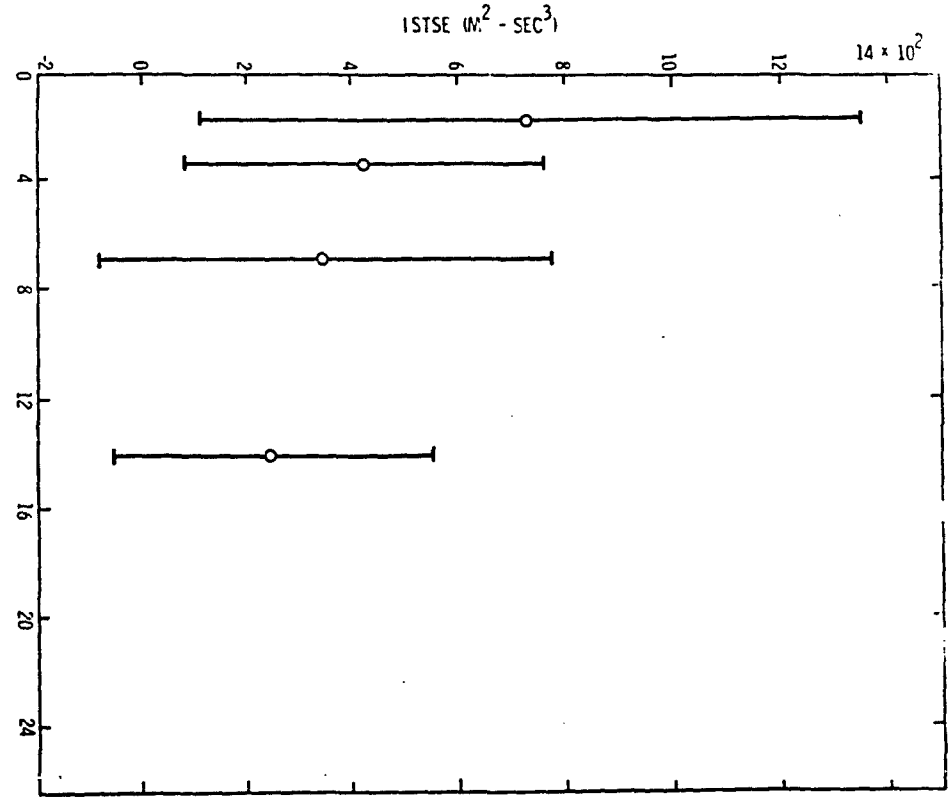


FIG. 8a EFFECT OF NATURAL FREQUENCY ON STEP GUST PERFORMANCE WITH LATERAL POSITION AS THE ERROR

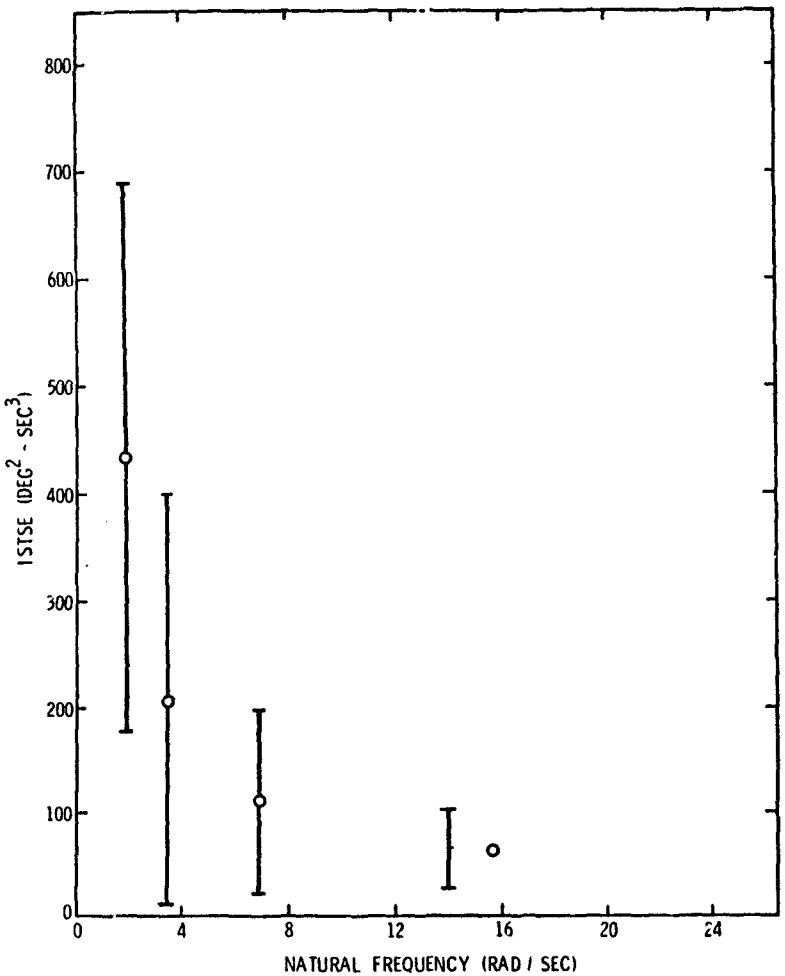


FIG. 8b EFFECT OF NATURAL FREQUENCY ON STEP GUST PERFORMANCE WITH YAW ANGLE AS THE ERROR

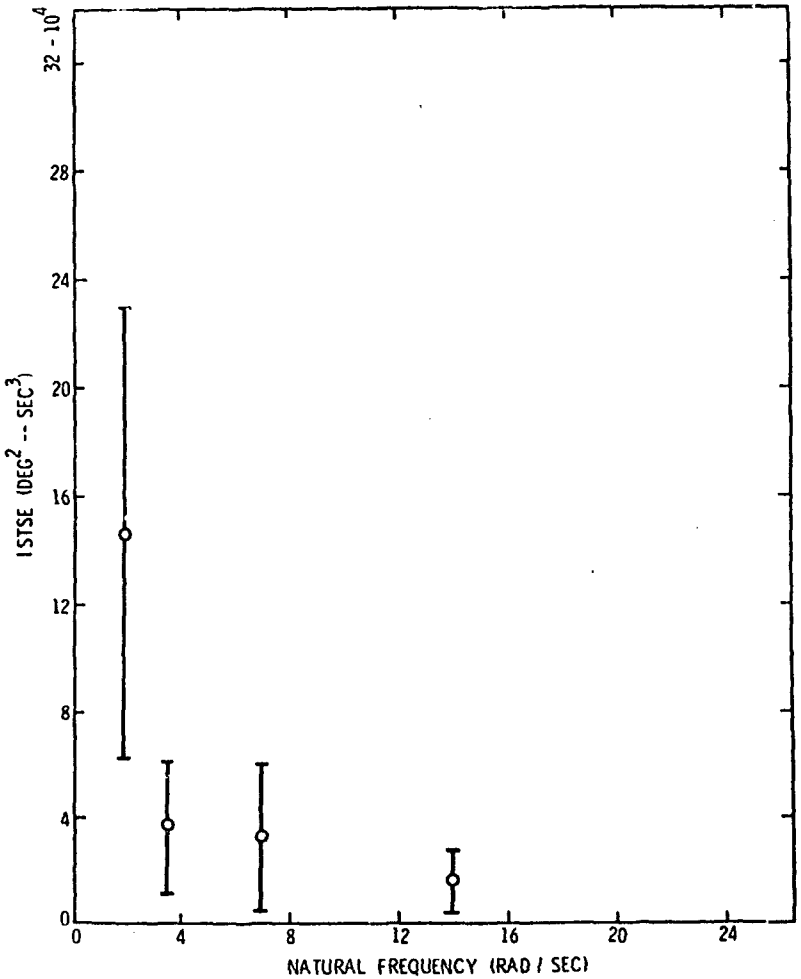


FIG. 8c EFFECT OF NATURAL FREQUENCY ON STEP GUST PERFORMANCE WITH STEERING WHEEL ANGLE AS THE ERROR

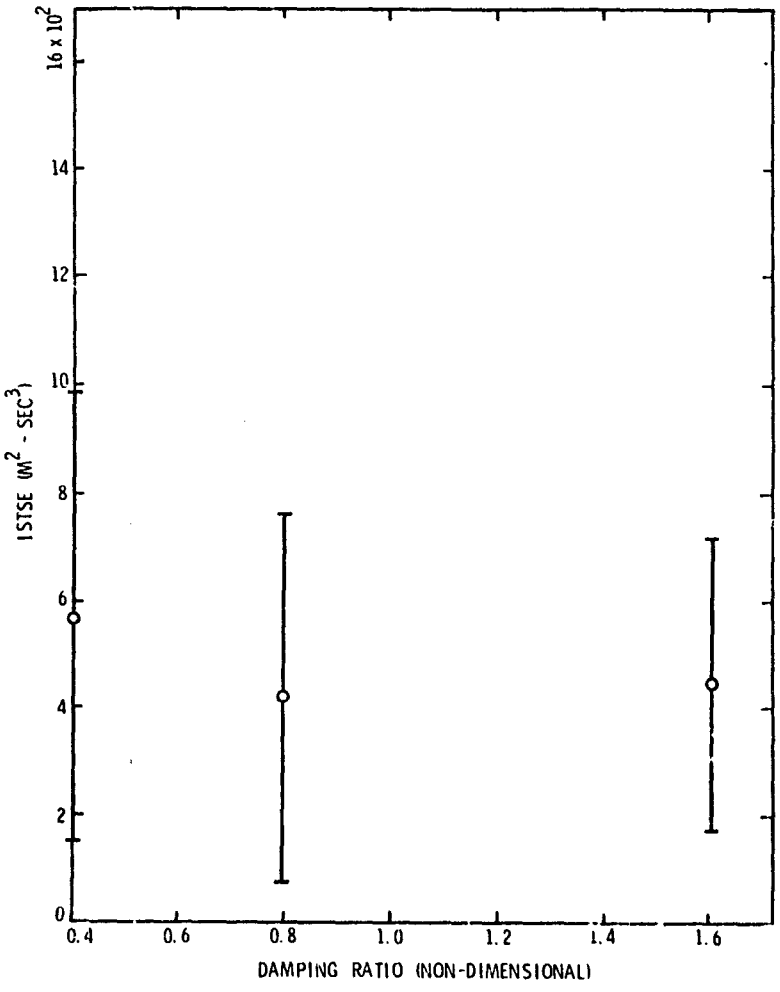


FIG. 9a EFFECT OF DAMPING RATIO ON STEP GUST PERFORMANCE WITH LATERAL POSITION AS THE ERROR

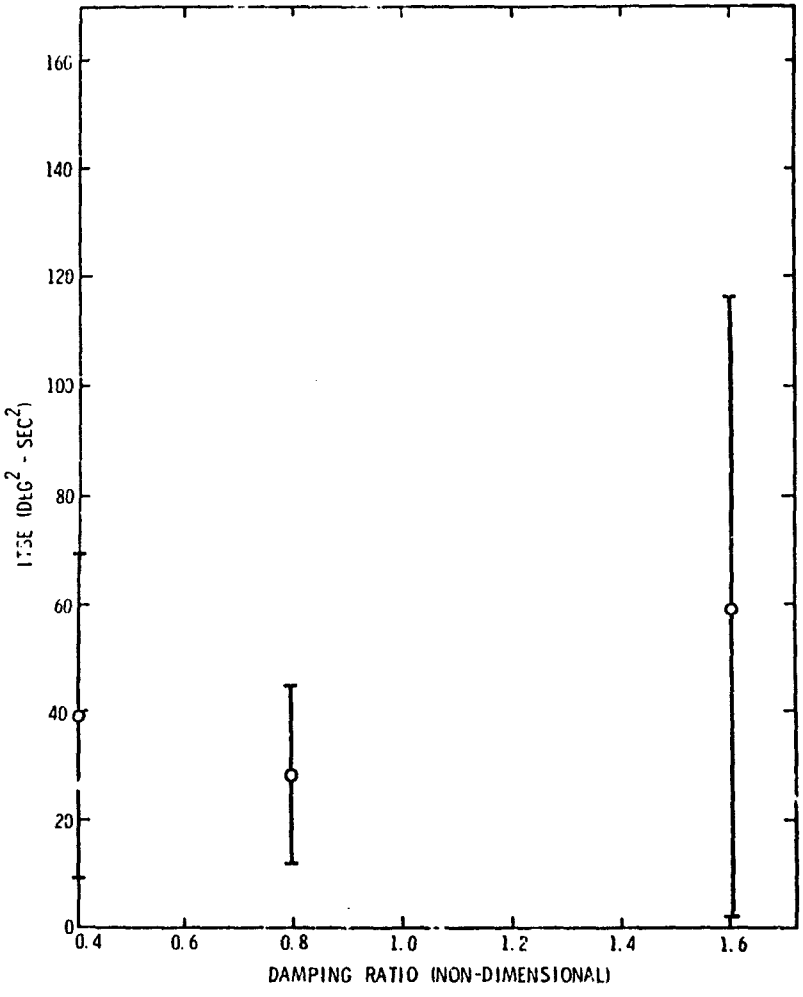


FIG. 9b EFFECT OF DAMPING RATIO ON STEP GUST PERFORMANCE WITH YAW ANGLE AS THE ERROR

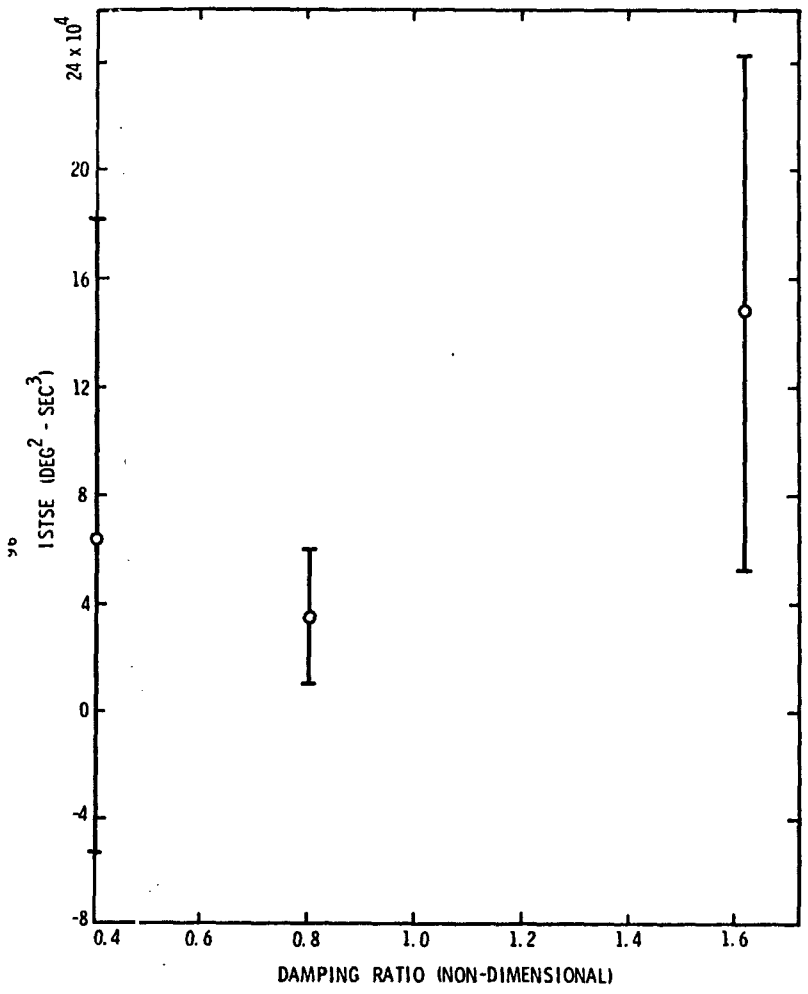


FIG. 9c EFFECT OF DAMPING RATIO ON STEP GUST PERFORMANCE WITH STEERING WHEEL ANGLE AS THE ERROR

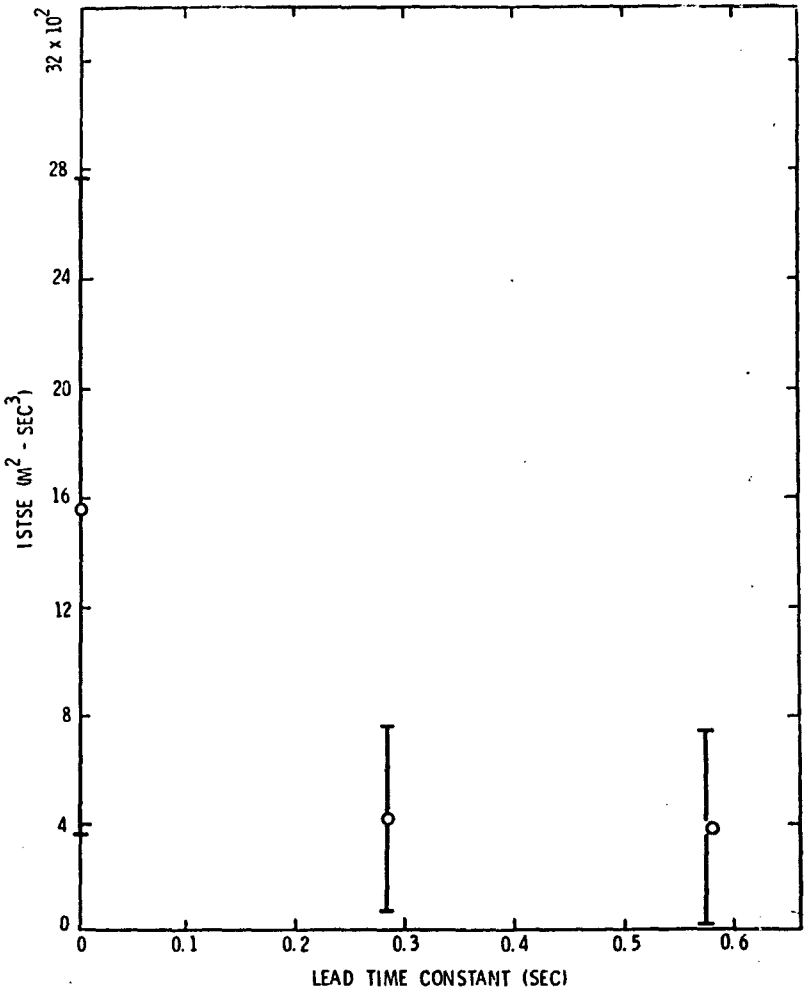


FIG. 10a EFFECT OF LEAD TIME CONSTANT ON STEP GUST PERFORMANCE WITH LATERAL POSITION AS THE ERROR

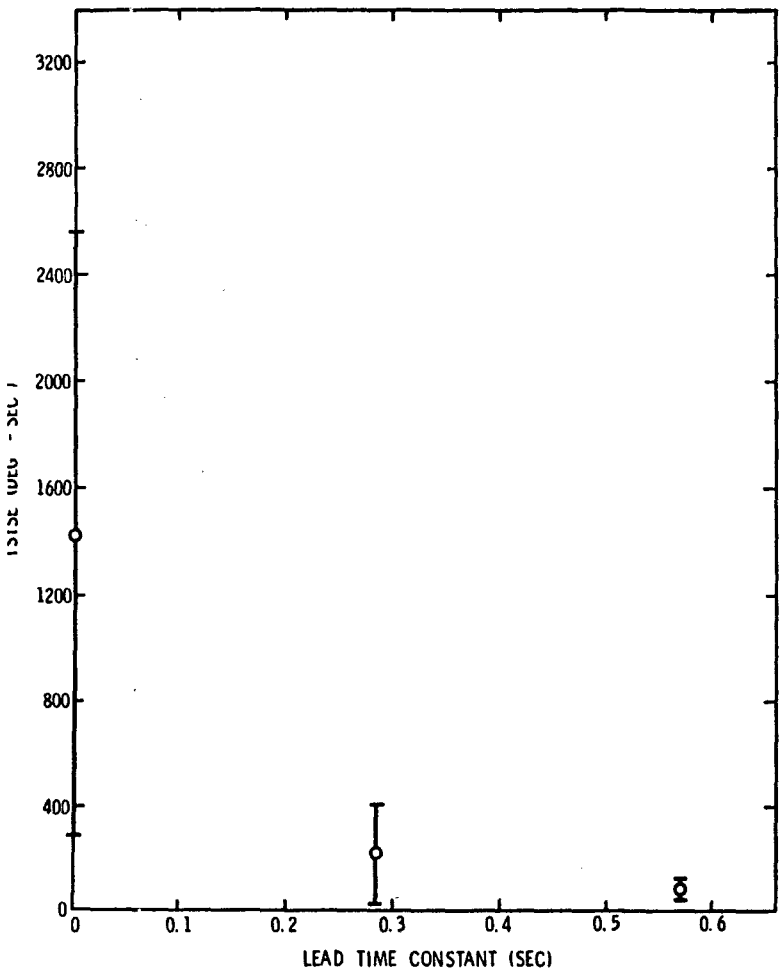


FIG. 10b EFFECT OF LEAD TIME CONSTANT ON STEP GUST PERFORMANCE WITH YAW ANGLE AS THE ERROR

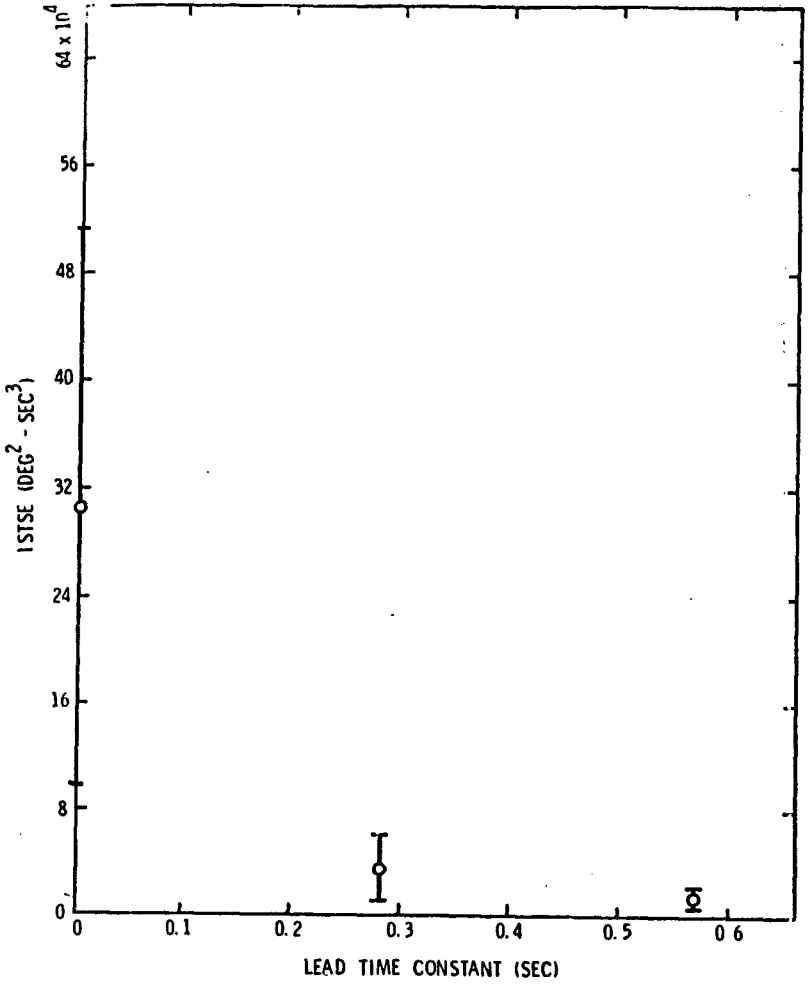


FIG. 10c EFFECT OF LEAD TIME CONSTANT ON STEP GUST PERFORMANCE WITH STEERING WHEEL ANGLE AS THE ERROR

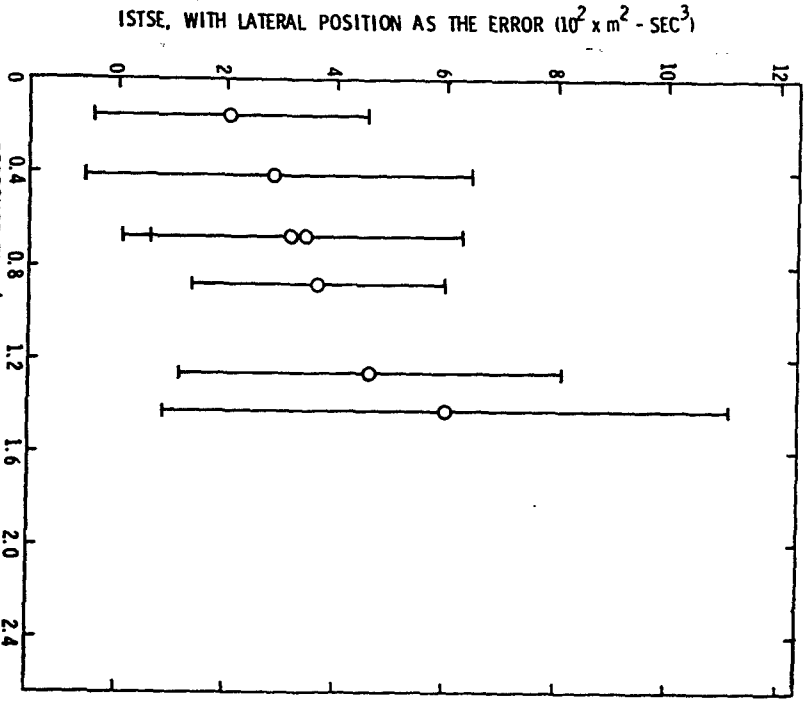


FIG. 11 PERFORMANCE DURING STEP GUST REGULATION TASK, WITH RESPONSE TIME DIVIDED BY EFFECTIVE DAMPING AS THE INDEPENDENT VARIABLE

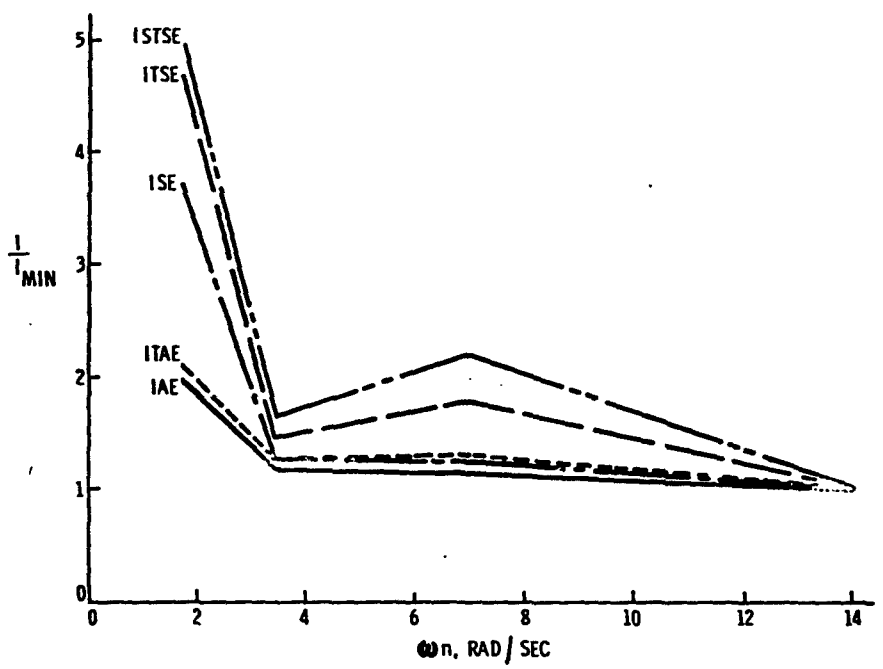


FIG. 12 COMPARISON OF PERFORMANCE CRITERIA AS A FUNCTION OF NATURAL FREQUENCY WITH STEERING WHEEL ANGLE AS THE ERROR (ARTIFICIAL LATERAL POSITION ERROR)

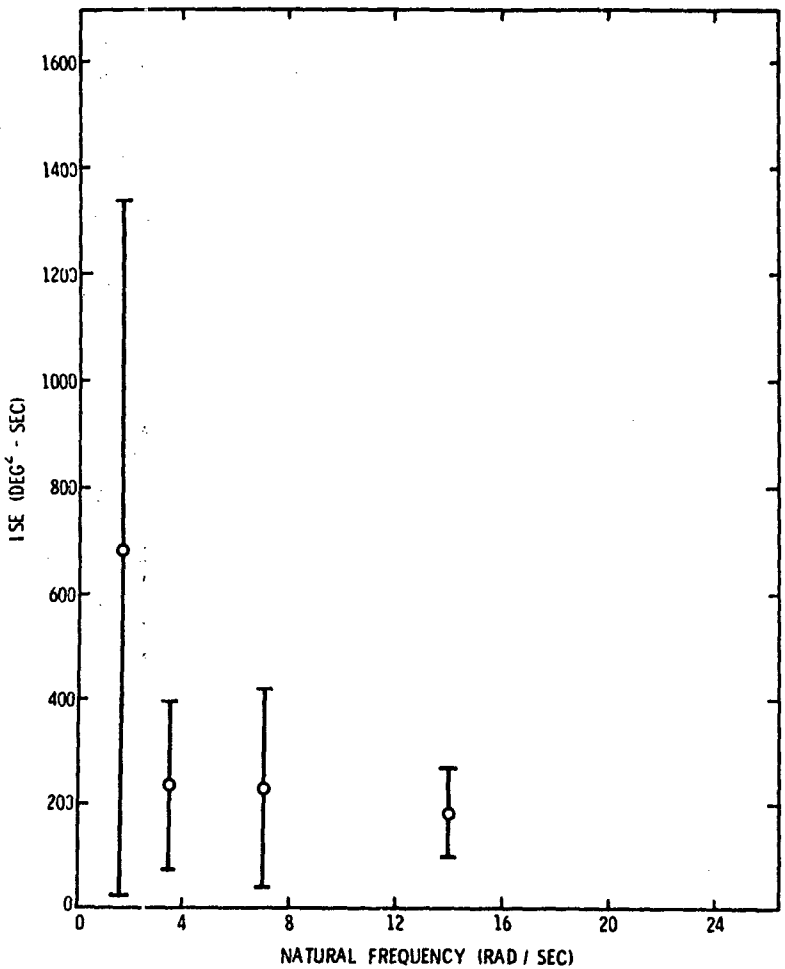


FIG. 13a EFFECT OF NATURAL FREQUENCY ON STEERING WHEEL BEHAVIOR FOLLOWING A STEP CHANGE IN LANE POSITION

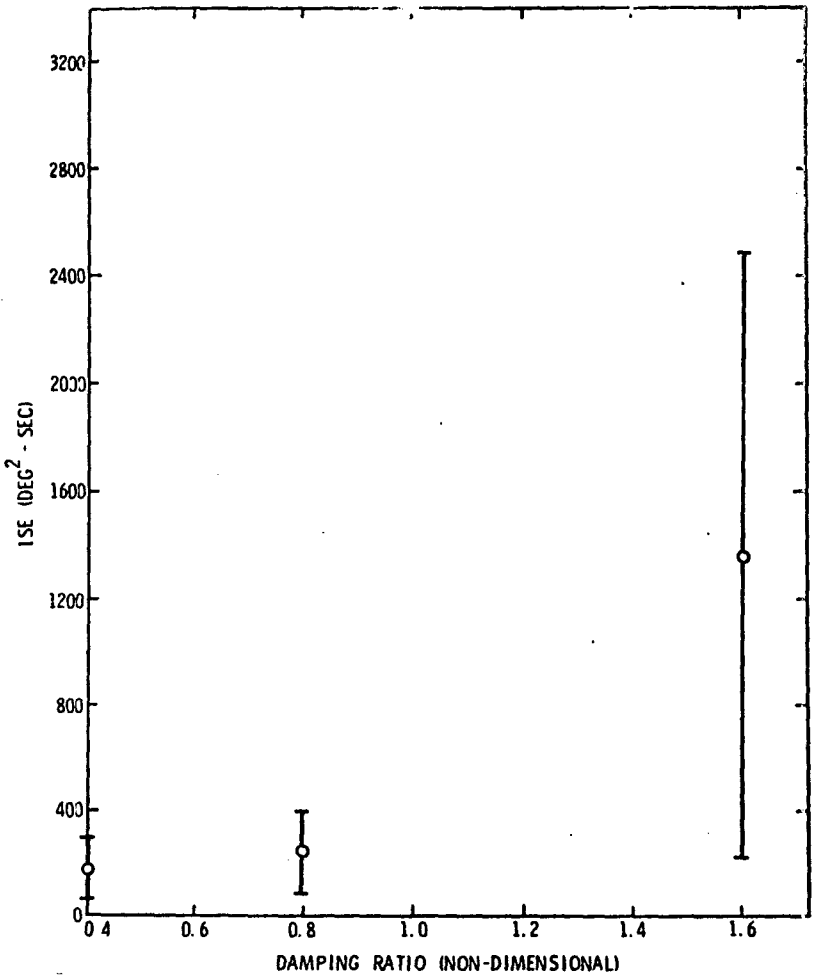


FIG. 13b EFFECT OF DAMPING RATIO ON STEERING WHEEL BEHAVIOR FOLLOWING A STEP CHANGE IN LANE POSITION

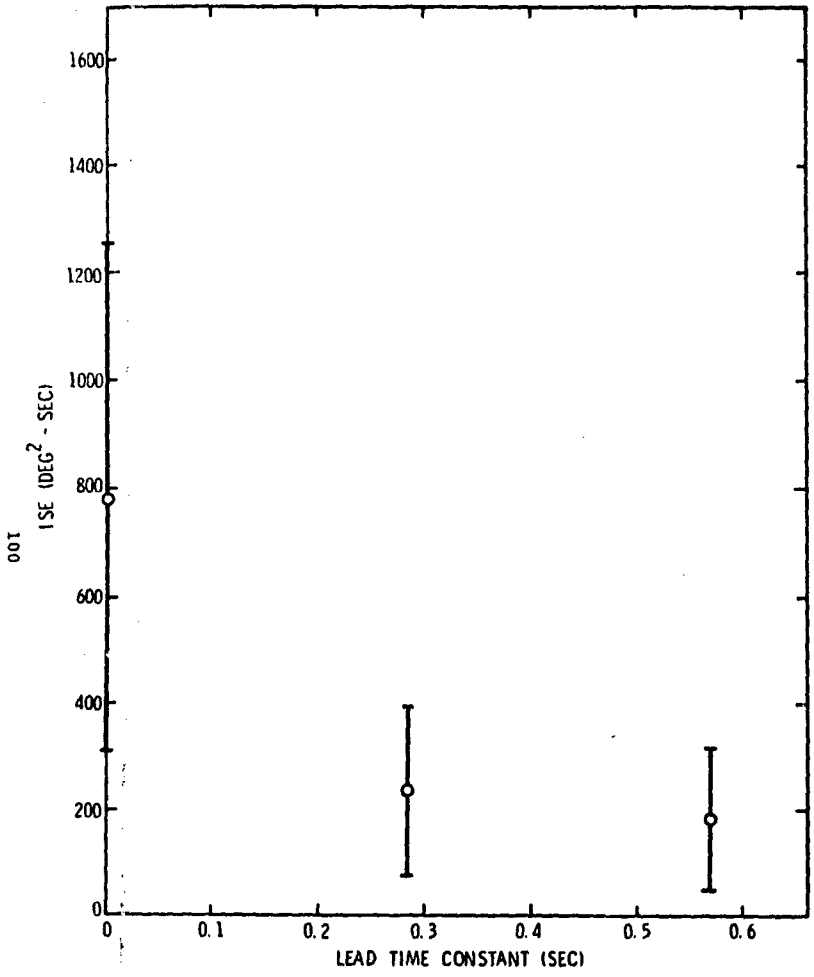


FIG. 13 C EFFECT OF LEAD TIME CONSTANT ON STEERING WHEEL BEHAVIOR FOLLOWING A STEP CHANGE IN LANE POSITION

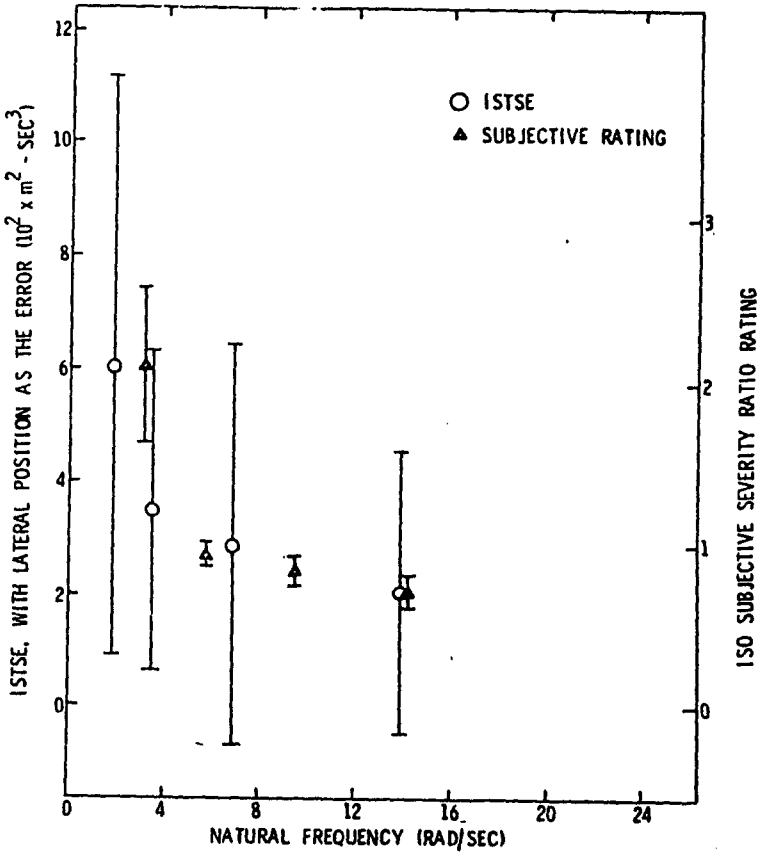


FIG. 14 EFFECT OF NATURAL FREQUENCY ON PERFORMANCE AND OPINION RATINGS DURING TRANSIENT MANEUVERS

THE FACILITATING EFFECTS OF UNCERTAINTY IN LONG-TERM MANUAL CONTROL*

William L. Verplank
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

SUMMARY

A 40-minute tracking task with different disturbance inputs has been used to look for the effects of reduced task demands on long-term manual control. The expected facilitating effects of task difficulty are hard to find. The decrements in performance over the run are no greater for the easier tasks. The detrimental effects of lower demand appear to be increased relative variability in performance, and possibly reduced performance on transition to unexpected, more difficult tasks.

An information measure, including the effects of "self-induced" uncertainty is developed as a work-load measure. There is a positive correlation between this "self-induced work-load" and performance decrement for the easiest task - just the opposite of what the facilitation hypothesis would predict.

THE FACILITATION HYPOTHESIS

Most people would agree that with increased automation there is a danger that the tasks left to the human operator might not be demanding enough; that somehow, a certain amount of task difficulty is required to facilitate human performance. These effects are expected to show up, if not over the short run, at least in long-term, low-complexity tasks.

The notion that a certain amount of stress is good is known variously as the "activation-" or "facilitation-hypothesis" or the Yerkes-Dodson Law.[1] It is usually represented as an inverted-U relationship between performance and "arousal". For this paper we use tracking error as the performance measure so the facilitation hypothesis would be represented as in Figure 1. (Note that there is expected to be an optimum stress between "boring" and "fatiguing".)

*This paper will be presented at the IEEE-SMC Conference, Sept. 1977.

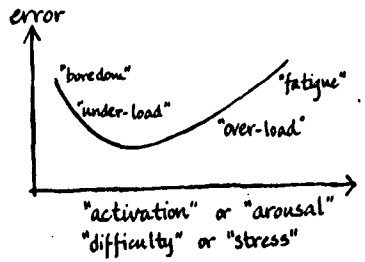


Figure 1. The facilitation hypothesis.

Monitoring and signal-detection tasks of low demand have been studied extensively within the vigilance paradigm.[2,3] Even though decrements in tracking performance follow the same trend as do decrements in detection performance, models of vigilance have not been applied to tracking and no measures of task difficulty have been developed which predict the effects of "under-load". Surprisingly, the literature contains little empirical support for the facilitation hypothesis.[4]

AN EXPERIMENT

To explore the effects of task difficulty on long-term manual control, subjects tracked for 40 minutes with one of four levels of difficulty. The dynamics were a double integrator ($1/s^2$); the control was a spring-centered stick. The display was a computer-generated image on a CRT that presented position and rate (heading) with a "perspective roadway".

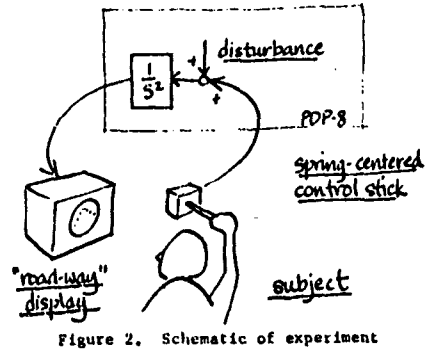


Figure 2. Schematic of experiment

Four conditions of disturbance ("task difficulty") were used (A,B,C,D), including no-input (A). Each disturbance consisted of the sum of 9 sine waves added to the human operator's (h.o.) output. The task simulated driving down a straight road with random lateral wind gusts.

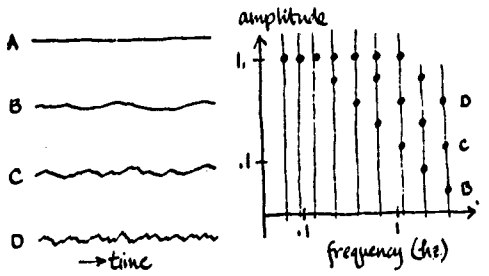
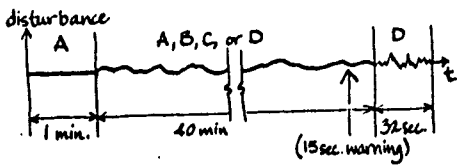


Figure 3. Disturbance samples and spectra.

The same disturbance was given for 40 minutes at which time the most difficulty disturbance (D) was given for the final 32 seconds. Each of the four conditions was given twice: once with and once without a warning before the difficult disturbance at the end. The warning consisted of three dots appearing in the road and moving "toward" the screen. They were visible for 15 seconds before they moved off the bottom of the screen. Three subjects completed 8 sessions each.



session*	1	2	3	4	5	6	7	8
disturbance	D	C	B	A	A	B	C	D
warning ?	yes				no			

*same order for all three subjects.

Figure 4. Experimental conditions.

Results: Long-term Performance Slope

Absolute error (from road centerline) was averaged every 32 seconds (E). Performance changes over the 40-minute run were taken as the average slope of E as a function of time, calculated with a least-squares linear regression. In 12 of the 24 runs, the slopes were significantly positive (indicating a performance decrement); in one case there was a significant, negative slope (improvement). ($n=72$, $\alpha=0.05$, $t > 2.0$ for significance)

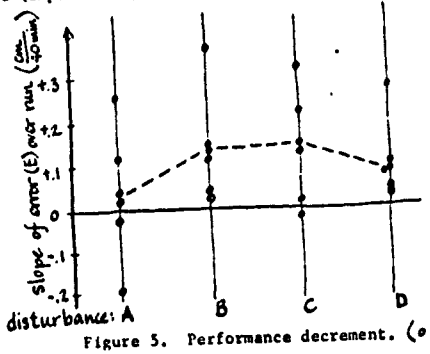


Figure 5. Performance decrement. (α - N.S.)

Taken together, there is no significant difference as a function disturbance but the trend seems to be in the direction opposite to what the facilitation hypothesis would suggest. There does, however, appear to be greater variability among the runs in the no-input case (A). Relative variability within runs was calculated as the ratio of the standard deviation of E (σ_E) to its mean ($\bar{E}$). With this measure, the no-input case (A) is worse, as might be suggested by the facilitation hypothesis. This is the same result observed in an earlier experiment using a CCTV driving simulator.[5]

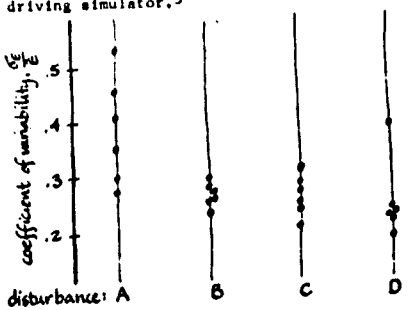


Figure 6. Relative variability within runs.

Results: Transition to Difficult Task

The last 32 seconds of each run was condition D. Performance as a function of the pre-condition (Figure 7) shows the same sort of results as for the average decrements (Figure 5): no significant difference between conditions and greater variability for the low-difficulty condition (A: no input). But the trend in this case is as the facilitation hypothesis would predict: better performance at intermediate difficulties.

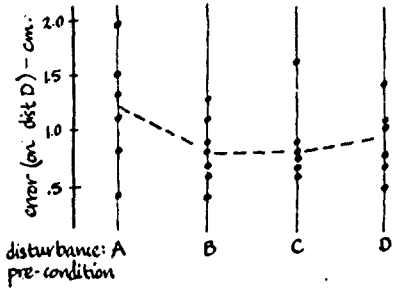


Figure 7. Performance (on D) at end of run.

INTERNAL UNCERTAINTY AS FACILITATION

Of particular interest is the case where there is no external disturbance. In some of these runs performance deteriorated, in some it did not and in at least one case it improved. Is there some way of predicting from a subject's instantaneous performance whether this is going to be a run where performance deteriorates or improves? One hypothesis is that the h.o. can "make the task interesting" or facilitating and thus sustain performance and avoid decrements.

Parameter Identification

The measure proposed in the next section requires identification of the h.c. transfer function. This was accomplished by identifying K_1, K_2 and τ in the following model from records of the h.o. output.

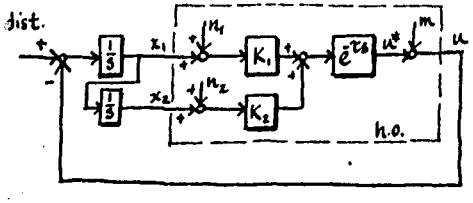


Figure 8. Parametric model used for identification.

Parameters K_1 and K_2 were estimated for different assumed τ by multiple regression $u(t) = K_1 x_1(t-\tau) + K_2 x_2(t-\tau)$, $t = nT$, $T = .03125$, $n = 1,2,\ldots,4096$. (ie 2 min at 32 Hz). Then the τ that gave the highest multiple correlation was chosen along with its K_1 and K_2. This method worked well, always giving a clear maximum correlation for τ in the expected range (.25 to .50 sec). The method of assuming a time-delay gives unbiased estimates, even without an external input disturbance, if the autocorrelation of the remnant is zero beyond τ.[6]

A Model for Internal Uncertainty

The measure proposed here as an indicator of facilitation is the information transmission rate. With no external disturbance input, the only information being transmitted is generated by the h.o. Two noise sources are hypothesized:

1. "input uncertainty" (n_1, n_2)
   band-limited white noises added to each observation and with variance proportional to the power of the variable being observed ($\sigma_{n_i}^2 = k_n \sigma_{x_i}^2$), and

2. "output uncertainty" (m)
   a white noise (limited to the same band) added to the h.o. output with $\sigma_m^2 = k_m \sigma_u^2$.

The information model used is for independent additive Gaussian signal (power density S(f)) and noise (power density N(f)).

$$I = \int_0^\infty \log_2 \left(1 + \frac{S(f)}{N(f)}\right) df \qquad (1)$$

In my case, I assumed that what is being transmitted is the operator's uncertainty in his own output (m), the corrupting noise is due to n_1 and n_2, and the output of the channel is the "intended output" u^*. That is, the signal-to-noise ratio is computed at u^*.

Noise ratio calibration. Without direct access to these hypothesized noise processes, constant noise ratios (k_n, k_m) were assumed. For input uncertainty $k_n = 0.01$ was used based on results from the literature on manual control[7] and with corroboration from psychophysics[8].

No direct measurements have been made of motor noise ratios, so a simple calibration experiment was performed. The subjects were told to (without looking) make repeated, equal-amplitude, back-and-forth motions of the control-stick in time with a metronome. The same three subjects used four rates (1,2,3,4 moves/sec) and four approximate target distances (.5, 1, 2, 3 cm). The mean and standard deviation of movement distance were calculated for sequences of 20 moves where the extremes of motion were taken as the beginning and end of successive moves.

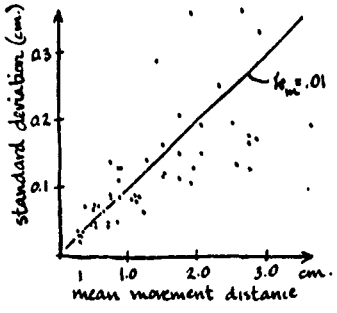


Figure 9. Output uncertainty from "blind-tapping" task.

The results show a constant ratio of the standard deviation to the mean (0.1) which is the same for all subjects and independent of frequency. Thus for output uncertainty, $k_m = 0.01$ which is the same ratio as for input uncertainty.

(In the information calculation k_n and k_m appear only as the ratio k_n/k_m, which is here assumed constant; a different value would shift all the results in the same direction: relatively more output uncertainty makes for more information transmitted.)

The "signal" and "noise" spectra at u^* are determined by the closed-loop transfer functions u^*/m, u^*/n_1, and u^*/n_2, which are calculated from parameters K_1, K_2 and τ taken from the h.o. identification.

$$\frac{S(f)}{N(f)} = \frac{\left|\frac{u^*}{m}\right|^2 k_m \sigma_u^2}{\left|\frac{u^*}{n_1}\right|^2 k_m \sigma_{n_1}^2 + \left|\frac{u^*}{n_2}\right|^2 k_n \sigma_{n_2}^2} \qquad (2)$$

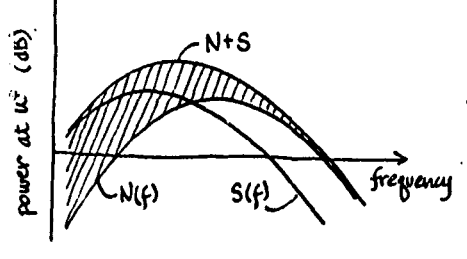


Figure 10. Components of power at u^*.

The shaded area in Figure 10 represents the integral of log (S+N) - log (N), and thus the information transmitted, (equation (1)).

Substituting the parametric model for the transfer functions gives a formula for the signal-to-noise ratio as a function of frequency. The identified parameters K_1 and K_2 only appear as a ratio $T_L = K_1/K_2$ corresponding to the "lead time-constant" and the time-delay, τ, does not appear.

(3)

T_L, $\sigma_{K_1}^2$, $\sigma_{n_2}^2$ and σ_u^2 were measured for each two-minute segment of each run and then used in the model (1) and (3) to calculate the information transmitted (I). The averages of I for the six runs with no input disturbance show a positive correlation with the decrement in performance over the run (Figure 11).

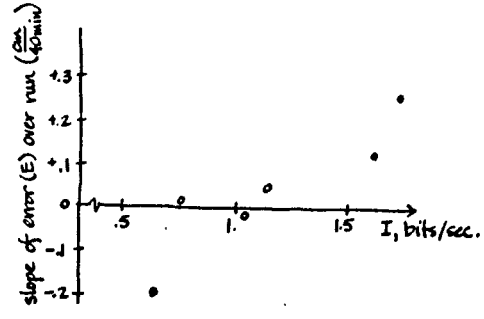


Figure 11. Performance decrement vs. "internal" information transmitted with no-input disturbance (A).

This is opposite to what the facilitation hypothesis would predict for low-stress tasks. It was thought that more information transmission or "self-induced work-load" would produce less performance decrement.

CONJECTURES

There are several possibilities: the proposed measure may not be a good indicator of facilitation, performance decrements might not be sensitive to facilitation, the facilitation hypothesis might be wrong in this case.

We do know that using just the external disturbance as a measure of difficulty (A,B,C,D) it is more difficult to predict what is going to happen over the run for the easier tasks and that some of that variability is correlated with the amount of "self-induced work-load".

REFERENCES

1. Kahneman, D., Attention and Effort, N.J.: Prentice-Hall Inc., 1973.

2. Mackworth, J.F., Vigilance and Habituation, England: Penguin Books Ltd., 1969.

3. Mackworth, J.F., Vigilance and Attention, England: Penguin Books Ltd., 1970.

4. Wiener, E.L., "On simultaneous monitoring and t cking", Journal of Applied Psychology v.60(1) (Feb., 1975), pp.100-105.

5. Verplank, W.L., "Is there an optimum work-load in manual control?", 12th Annual Conf. on Manual Control, NASA TM X-73, 170, 1976.

6. Wingrove, R.C., "Comparison of methods for identifying pilot describing functions from closed-loop operating records", NASA TN D-6235, 1971.

7. Levison, W.H., Baron, S., Kleinman, D.L., "A model for human controller remnant", 5th Ann. Conf. on Manual Control, NASA SP-215, 1969.

8. Crossman, E.R.F.W., "Discussion of Paper 9: A Model for Human Controller Remnant by Levison, Baron, and Kleinman", NASA SP-215, 1969.

PERFORMANCE AND WORKLOAD ANALYSIS OF IN-FLIGHT HELICOPTER TASKS*)

by P.H. Wewerinke

National Aerospace Laboratory NLR

SUMMARY

The study described in this paper was aimed at assessing the potentials of the optimal control model structure to predict the important characteristics of realistic operational helicopter missions.

The theoretical and experimental results indicate that the optimal control model succesfully predicts the best attainable (rather than the average) performance of a group of well-trained, highly motivated subjects. Furthermore, the model allows a description of inter-subject variability.

The control effort model predictions have been supported by subjective ratings. The model seems to provide a meaningful representation of pilot workload involved in complex control tasks.

INTRODUCTION

Many questions related to operating aircraft are still largely unresolved. The available handling quality criteria predominantly reflect the effect of the pilot on the vehicle. By the same token, operational research studies often exclusively concentrate on system performance. In spite of an increasing awareness of the necessity to emphasize the human operator's participation in manned vehicle systems, there are still very few satisfactory tools to describe his role in these systems. A promising approach concerns the use of mathematical models of human behavior. Among several different approaches toward human operator modeling, the optimal control model has emerged as the most useful one for the study of complex manned aerospace systems.

This model has been tested, so far, only against data obtained for (a variety of) laboratory tasks. The objective of the present study was to investigate the usefulness of the optimal control model structure (Refs.1-3) and a control effort model (Ref. 7) to describe realistic, in-flight aircraft control tasks. Especially, the use of the model as a predictive tool was concentrated on. For this,three helicopter instrument control tasks were investigated: a hover and two navigation tasks at prescribed heights.

In the following chapter some aspects of the optimal control and workload model are discussed. Next, the experimental program is described and the experimental results are compared with model predictions

OPTIMAL CONTROL MODEL

The helicopter control tasks concerned are described in terms of the optimal control model (OCM). The inputs of the model which is documented extensively in the literature (e.g. Refs. 1 and 2) can be divided in task-related characteristics (parameters) and human operator parameters. The former comprise the system (helicopter) dynamics, the disturbance environment (moderate atmospheric translational gust is included) and the information available to the pilot to perform the task (via an integrated display). The latter can be considered as the real input parameters of the model or assumptions to be made:

- the objectives of the task (instructions) yielding a given control strategy. This is represented by the weighting matrices of the quadratic cost functional which is (assumed to be) minimized by the pilot's optimal control strategy
- an equivalent perceptual time delay which has been found to be relatively constant (0.20 sec)
- the attention dedicated to (observation noise corresponding with) the various displayed variables
- human randomness in executing control inputs which will be represented by a (constant) motor noise ratio of -25 dB (disturbances enter the system in parallel with the controls).

The cost functional weightings are selected on the basis of maximum allowable deviations, or limits. The choice of these for the present study will be discussed in the next chapter. The weightings for each quadratic term in the cost functional is then the inverse of the square of the corresponding limit.

The observation noise covariances have been found to be proportional to the mean-squared value of the displayed variables and inversely proportional to the square of the random input describing function gain associated with a threshold device. This threshold is in the following identified with the "acceptable" deviations of displayed variables ("indifference" thresholds). Furthermore, the interference model of reference 3 indicates that the covariances are inversely proportional to the fraction of attention f_i, paid to the displayed variable y_i. In formula

*) This investigation has been carried out under contract for the Research Branch of the Directorate of Material Air RNLAF.

$$V_{y_i}(t) = \frac{\tau P_{O_i} E\left\{ y_i^2(t) \right\}}{f_i K_{y_i}^2 (\sigma_{y_i}(t), a_{y_i}(t))} \qquad (1)$$

where P_{O_i} is the "noise/signal" ratio corresponding with "full attention" and has units of normalized power (positive frequencies) per rad/sec and K_{y_i} is the afore-mentioned gain.

For the display situation of the experimental program it is assumed that P_{O_i} is equal for all displayed variables: P_O. This indicates the overall level of attention paid to the task and is, just like the "indifference" threshold, a variable in the subsequent analysis. The allocation of attention among the displayed variables can be optimized (minimal cost functional). The formulation of this problem and its experimental support is contained in the next section.

Once the cost functional weightings, the "indifference" threshold and the overall level of attention are assumed (model inputs), a variety of measures of performance (variances are considered in the following) can be obtained as well as the corresponding workload. The latter is provided by the control effort model discussed in the following.

Optimal allocation of attention

In accordance with the fundamental assumption of optimality it can be assumed that the human operator divides his attention among the displayed variables optimally, i.e., minimizing the cost functional J (Ref. 4). So the problem is to minimize $J(f_y)$ subject to the (scalar) equality constraint

$$\sum_{i=1}^{M} f_{y_i} = 1 \qquad (2)$$

where M is the number of display indicators, and the (vector) inequality constraint

$$f_{y_i} \geq 0 \text{ for } i = 1,\ldots M$$

This is a standard parameter optimization problem which can be solved numerically by a first-order gradient method.

There is ample experimental evidence that the human operator derives from a given display indicator both displacement- and rate information*.

* It has been found that sometimes the human operator obtains also acceleration information from a display indicator (Ref. 6)

Relatively good model results have been obtained assuming equal attention (i.e. equal observation noise ratio) for displacement and rate. However, an alternative hypothesis is that the human operator divides his attention between displacement- and rate information. It is (again) assumed that this allocation of attention will be optimal (minimal cost functional) with the constraint $f_y + f_{\dot{y}} = 1$; in other words, the human operator divides his capacity optimally between all the display variables (y_1, $\dot{y}_1$, ... y_M, $\dot{y}_M$).

Excellent experimental support for the foregoing hypothesis is provided by the model results of the experimental program of reference 7. In this study a variety of single-indicator tracking tasks (controlled element dynamics) was investigated. Accurate model results were obtained without the constraint of equal observation noise ratio for error displacement and error rate. Position, velocity and acceleration control tasks (K, K/s and K/s^2 dynamics, respectively) were considered with various instability levels. The results of four typical configurations are shown in figure 1. The "measured"** allocation of attention between error displacement and error rate is clearly quite close to the optimum (minimal cost functional). Although for the K/s^2 tasks no reliable (unique) observation noise ratios could be obtained**, also for these configurations the experimental results suggest an allocation of attention near the theoretical optimum (a fraction of attention to error rate of about 0.9).

Based on the foregoing results it will be assumed in the following analysis that the human operator (model) divides his information processing capacity optimally among the display variables (displacement and rate of all display indicators).

Control effort model

For a complete description of human control behavior and its impact on overall system reliability, it is necessary to assess how hard the human controller has to work to achieve a given performance (criterion). Because of the adaptive capabilities of the human operator control effort is often the most sensitive to control task characteristics under consideration. In this context a control effort model is developed in terms of the optimal control model parameters. This model is presented in reference 7.

* By matching performance scores of interest, as well as describing function and remnant data, model parameters (observation noise ratios) could be estimated quite accurately resulting in the "measured" fractions of attention.

** The performance measures are very insensitive to position noise ratio, furthermore, because of the important effect of motor noise, no unique rate observation noise ratio could be obtained.

Since a somewhat modified version of the model is used in the subsequent analysis (directed at multivariable control situations), the model is briefly reviewed in the following.

Human control response is partly determined by mechanisms that selectively tune the organism to the stimulus situation, by which is meant both selectively attending to some stimulus in preference to others and investing more or less attention per source of information. This can be identified with voluntary attention (Ref. 8), reflecting that the subject attends to the stimulus because of its relevance for performing the task and not only because of its arousal function. Also involuntary attention is included in the control effort model. This can be related to the level of arousal and is largely dictated by the properties of the displayed information.

The aspect of voluntary attention is incorporated in the model in terms of the overall level of attention, P_O. The aspect of involuntary attention is included in the control effort model in terms of the sensitivity of task performance (cost functional, J) to the momentary attention paid by the subject. In formula

$$E = S/P_O \text{ (dB)} \qquad (3)$$

with

$$S = \frac{\partial J}{\partial P_O} \text{ (dB)}$$

where the partial derivative indicates that the other model parameters are kept constant.

For the eight single-axis control tasks of the experimental program presented in reference 7 the computed control effort results are compared with subjective ratings. The result which is shown in figure 2 exhibits an excellent correlation between both.

EXPERIMENTAL PROGRAM

Experimental conditions

Several considerations were involved in the experimental set-up. In order to define (i.e., measure and model) accurately the (in-flight) helicopter tasks, instrument tasks were chosen allowing a complete description of the displayed information. Another consideration was to include various configurations respresenting a sufficient variation in workload to obtain significant experimental results.

Based on this, an instrument hover task was chosen consisting of stabilizing an Alouette III helicopter at a height of 600 ft with minimal horizontal (ground) speed. The three attitude angles were provided by a three-axis ADI presented in figure 3. Horizontal velocity components were presented via the cross pointers shown in the figure. A backward velocity of 5 kts corresponded with a full needle deflection upwards (all the display signs were chosen according to the "fly-to" principle). A height error of 100 ft corresponded with a display deviation of 2 dots.

Furthermore, two navigation tasks were chosen consisting of flying along a desired track with an indicated airspeed of 60 kts at a prescribed height: 600 and 150 ft, respectively. A track deviation of 200 ft corresponded with a full deflection of the vertical pointer shown in figure 3. A height error deflection of 2 dots corresponded with a height error of 100 ft for the navigation task at 600 ft and of 60 ft for the task at 150 ft. The subjects were instructed to maintain a constant indicated airspeed of 60 kts provided by the display shown in figure 3. So, apart from this indicator, all the information to perform the task was provided by the ADI.

Each sortie consisted of two hover tasks (of 3 min) and the two navigation tasks (of 5 min). Each subject performed two training flights to partially eliminate learning effects. For several reasons the program (9 sorties per subject) could not be completed; this is indicated by the number of replications given in the following tables.

All helicopter parameters of interest were recorded digitally; furthermore, subjective ratings were collected on the rating scales given in table 1. The four participating helicopter pilots had (on the average) a flying experience of 1200 hours of which 60 hours instrument flying. Figure 4 shows the Alouette III helicopter with the evaluation pilot (seated to the right), the safety pilot and the observer.

Experimental and model results

Since space does not permit an extensive presentation of the experimental results (these are given in reference 9), only the principal experimental results will be discussed corresponding with the model results. This implies that the results of two subjects of whom sufficient experimental data were obtained will be emphasized in the following.

Model predictions

Optimal control model predictions were obtained on the basis of the following assumptions:

- the cost functional weightings selected via the maximum allowable limits were chosen on the basis of the available understanding of the task requirements and physical- and display limits. The result is shown in table 2
- the indifference thresholds were zero and the overall level of attention was obtained by determining the "optimum" trade-off between system performance and attention ("knee" of the curve) supported by the results of an experimental program (Ref. 10) using the same display.

The predicted system performance scores and the optimum allocation of attention (an equal division of attention between longitudinal and lateral control has been assumed) are presented in table 3 for the HOVER task and in table 4 for the NAV-H task. Also the corresponding measurements are given. In the tables an overall performance index, J_m, is shown. This performance index incorporates the scores which were instructed to be minimized: height error (h) and horizontal speed (v_h) for the HOVER task and height error and lateral deviation (y) for the NAV tasks. These scores are weighed by the corresponding display limits, so that J_m is the sum of the meansquared fractions of the full display deviations. This criterion is analogous to the cost functional of the optimal control model were it is assumed that this criterion is minimized by the human operator.

The results in table 3 indicate that there is a substantial difference in hover performance between the three subjects. The model predictions concerning the guidance variables (height error, h, and total horizontal speed, v_h) are clearly too optimistic. However, the trend in performance between the subjects leads to the conclusion that the model predictions reflect the "limit" (optimum) of human control behavior (of the well-trained, well-motivated pilot). This is also applicable to the results of the NAV-H task. For this task the inter-subject variability is considerably less than for the HOVER task; therefore, also the average performance is given in table 4.

The experimental results of the NAV-L task are only significantly different from the NAV-H results with respect to the height performance. The model did (exactly) predict this performance improvement for the NAV-L task (RMS h_L/RMS h_H = 0.8).

The assumption that no thresholds are involved in observing the display quantities is reasonable to the extent it is related to the quality of the displays involved in the experiment. This was the consideration for neglecting the thresholds. However, the assumption is not in accordance with pilot's control behavior in real flight: within certain limits the pilots tolerate display deviations (do not take any control actions) (Refs. 11 and 12). Thus, a second "prediction" was made assuming an indifference threshold (TH) of 1/6 of the full display deflections of the guidance variables (h,u, v,y). Only the performance scores of the pertinent guidance variables were affected by this assumption (as might be expected). The resulting scores are given (between parentheses) in tables 3 and 4.

In summary, it can be concluded from the foregoing that the optimal control model predictions do reflect optimal control behavior, i.e., the model results represent the best attainable performance. To put another way, the model predictions may not be considered as describing the average pilot's control behavior but the best pilot's performance.

The following paragraph deals with the question which model assumptions have to be modified in order to obtain a better agreement between model and experimental results.

Model match

Based on the experimental and model results the following parameters were changed in order to match the experimental results of subjects A and B.

- The maximum allowable limits corresponding with the cost functional weightings on attitude angles and control deflections were somewhat diminished. The pertinent values are given in table 2.
- For subject A an indifference threshold of 1/6 was assumed (a threshold of 1/6 of the full display deflections of the guidance variables) and for subject B (considerably less motivated) a ratio of 1/? was taken.
- An optimal division of attention between longitudinal and lateral control was determined and the overall level of attention, P_o, was obtained by matching the overall performance index, J_m.

The resulting model scores are compared with the measured scores for both subjects for the HOVER task in table 5. Apart from the substantial difference in heading scores*, all the important performance scores match well**.

---

* This difference is possibly due to an inaccuracy in the description of the vehicle dynamics (stability and control derivatives). At any rate, even a substantial variation of all model parameters could not remove this discrepancy.

** In order to match the height score of subject B it had to be assumed that somewhat less attention was paid to height and forward speed information than the optimal amount.

Table 6 presents the model and experimental results of subject A of the NAV tasks. As the difference in performance between the NAV-H and the NAV-L task seems insignificant, no attempt has been made to model the tasks separately. All scores match quite well, suggesting that the model provides a good description of pilot behavior for both tasks. In table 7 the results of subject B are given for both NAV tasks. The NAV-H results were obtained as indicated before. For the NAV-L match the maximum allowable limit of 60 ft full display deflection was used for the weighting on height error[*]. It had to be assumed that subject B spent somewhat more attention to longitudinal control than subject A (.08). Again, all the model scores match reasonably well the corresponding measurements.

In summary, it can be concluded from the foregoing that for all the configurations considered a good agreement between measured scores and model results could be obtained on the basis of reasonable assumptions and basically two model parameters: the indifference threshold ratio and the overall level of attention. The first parameter can be related to the motivation of the pilot which appears to vary substantially between the subjects participating in the present experiment. The latter parameter reflects also to some extent the motivation of the subject. Moreover, it is partly dictated by the task (demand of the task). This will be discussed in the next section.

Control effort model results and subjective ratings

In this chapter the foregoing model results are used to compute the corresponding control effort. These results are compared with subjective ratings. Because of the limited data base this analysis must be considered as exploratory.

For subject A control effort has been computed for the HOVER task and the NAV tasks. Also the theoretical curves of system performance versus control effort have been established. The result is shown in figure 5. The model predicts that the HOVER task is more demanding than the NAV-tasks. Furthermore the figure shows that the HOVER performance is more sensitive to attention (or effort) than the NAV performance. This explains why the inter-subject variability is much larger for the HOVER task than for the NAV tasks and visualizes the greater demand of the HOVER task (the HOVER task "forces", but also enables, the subjects to spent that much effort). For subject B control effort has been computed for all three tasks. Again the model predicts that the highest workload is involved in performing the HOVER task.

The control effort model results are compared with subjective ratings in table 8. Also the ratings indicate that the HOVER task is the most difficult. Furthermore, on the average, there is no significant difference in effort between the NAV-L and the NAV-H task. This is even more clearly illustrated comparing the average ratings of all subjects with the corresponding model predictions. This is also presented in table 8.

In summary, it can be concluded that the control effort model predictions have been supported by the subjective ratings. This provides additional validation for the model which, so far, has been tested only against data obtained for (a variety of) laboratory tasks.

[*] This subject performed the HOVER task with various horizontal speed display sensitivities (5 and 25 kts full display deflections). He maintained a surprisingly constant level of horizontal speed performance in terms of display deviations and not in units of kts. This excellently supports the assumption involved in the optimal control modeling that the pilot's control strategy is such that the display deviations are within an "acceptable" region.

CONCLUSIONS

The objective of the present program was to investigate the usefulness of the optimal control structure and the control effort model to predict and describe performance and effort of realistic helicopter control tasks. Referring to the results presented in the previous chapter the following specific conclusions can be drawn.

- The optimal control model can succesfully be used to predict the (best attainable) performance of in-flight (helicopter) control tasks.
- In real flight circumstances the pilot tolerates display deviations within certain limits. This can be taken into account in the optimal control model by assuming (statistical) thresholds in perceiving these variables.
- The rationale for selecting the weightings in the cost functional of the model - which can be related to the assumption concerning the pilot's control strategy - is convincingly supported by the experimental results.
- The model provides a suitable framework to formulate differences in control behavior between subjects, basically in terms of two model parameters: the indifference threshold ratio and the overall level of attention. Both reflect personality traits, such as motivation.
- The control effort model predictions have been supported by subjective ratings for the three helicopter tasks under investigation. Although additional experimental support for the model is desirable, it seems to provide a meaningful representation of pilot workload involved in complex control tasks.

REFERENCES

1. Baron, S. et al.: Application of optimal control theory to the prediction of human performance in a complex task. AFFDL-TR-69-81, March 1970.

2. Kleinman, D.L. and Baron, S.: Manned vehicle system analysis by means of modern control theory. NASA CR-1753, June 1971.

3. Levison, W.H. et al.: Studies of multivariable manual control systems. A model for task interference. NASA CR-1746, May 1971.

4. Curry, R.E. et al.: A model for simultaneous monitoring and control. NASA TMX-62,464, May 1975.

5. Levison, W.H. et al.: Modeling the effects of environmental factors on human control and information processing. AMRL-TR-76-74, August 1976.

6. Levison, W.H.: The effects of display gain and signal bandwidth on human controller remnant. AMRL-TR-70-93, March 1971.

7. Wewerinke, P.H.: Effort involved in single- and two-axis manual control systems. NLR TR 75060 U, November 1974.

8. Kahneman, D.: Attention and effort. Prentice-Hall Inc. 1973.

9. Wewerinke, P.H.: An analysis of in-flight helicopter pilot control behaviour and workload. NLR TR 76146 C, December 1976.

10. Wewerinke, P.H.: Human control and monitoring - Models and experiments. NASA TM X-73, 170, May 1976.

11. Baron, S. and Levison, W.H.: A manual control theory analysis of vertical situation displays for STOL aircraft. BBN Rep. No. 2484, April 1973.

12. Ault, A.: Pilot workload analysis. Paper presented at the Monitoring Behaviour and Supervisory Control Symposium, Berchtesgaden, F.R. Germany, March 8-12, 1976.

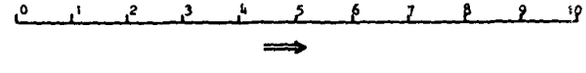
13. Bryson, A.E. and Ho, Y.C.: Applied optimal control. Blaisdell Publishing Company, 1969.

Table 1: Rating scales

Name:

Task:

Using the scale below, indicate the degree of effort you spend on performing the task



Rating scale for Controllability and Precision		Rating Scale for Demands on Pilot	
0	Extremely easy to control with excellent precision	0	
1		1	
2	Very easy to control with good precision	2	Completely undemanding, very relaxed and comfortable
3		3	Largely undemanding relaxed
4	Easy to control with fair precision	4	
5		5	
6	Controllable with somewhat inadequate precision	6	Mildly demanding of pilot attention, skill, or effort
7	Controllable, but only very imprecisely	7	Demanding of pilot attention, skill, or effort
	Difficult to control		
8	Very difficult to control	8	Very demanding of pilot attention, skill, or effort
9	Nearly uncontrollable	9	Completely demanding of pilot attention, skill, or effort
			Nearly uncontrollable
10	Uncontrollable	10	Uncontrollable
	Not applicable		Not applicable

Parameter		HOVER allowable limit	NAV allowable limit
θ	(rad)	0.15 (0.1)	0.1
ϕ	(rad)	0.15 (0.1)	0.1
ψ	(rad)	0.15 (0.1)	0.1
u	(kts)	5	10 (15)
v	(kts)	5	-
h	(ft)	100	100
y	(ft)	-	100(H) - 60(L)
δ_e	(rad)	0.2 (0.1)	
δ_a	(rad)		
δ_r	(rad)		
CP	(rad)		
$\dot{\delta}_e$	(rad/sec)	0.4 (0.2)	
$\dot{\delta}_a$	(rad/sec)		
$\dot{\delta}_r$	(rad/sec)		
CP	(rad/sec)		

between parentheses the values used for the matched model results; otherwise, the pre- and post-experimental values are identical.

Table 2: Allowable limits used for the cost functional weightings

PARAMETER	MODEL PRED.	MEASURED SUBJECT A	B	C
σ_θ (deg)	1.6	1.7	1.6	1.4
σ_ϕ (deg)	1.0	1.6	1.3	1.6
RMS ϕ (deg)	4.3	2.5	4.8	4.5
RMS h (ft)	13.1(16.3)	21.3	55.1	100.4
RMS u (kts)	0.8(1.0)	1.5	2.5	2.8
RMS v (kts)	0.4(0.7)	1.2	1.6	2.2
RMS v_h (kts)	0.9(1.2)	1.9	3.0	3.5
σ_{δ_e} (deg)	1.1	1.1	1.0	0.7
σ_{CP} (deg)	1.1	1.1	0.8	0.4
σ_{δ_a} (deg)	0.2	0.6	0.5	0.6
σ_{δ_r} (deg)	1.7	1.1	0.6	0.7
Overall performance J_m	0.05(0.09)	0.20	0.66	1.49
Replications	-	15	5	3

(.): predictions with thresholds

ALLOCATION OF ATTENTION	
θ	.00
$\dot{\theta}$	.11
ϕ	.00
$\dot{\phi}$	.01
$\phi \cdot \phi$	.05
$\dot{\phi} \cdot \phi$	.14
h	.05
$\dot{h}$	.04
u	.02
$\dot{u} \cdot u$	.29
v	.02
$\dot{v}$	.28
τ_{LOHC}	.5
P_O	-18 dB

Table 3: Model predictions and experimental results for the HOVER task

PARAMETER		MODEL PRED.	MEASURED SUBJECT A	B	C	D	AVERAGE
σ_θ	(deg)	0.9	2.6	1.9	2.2	1.8	2.2
σ_φ	(deg)	2.8	2.8	2.4	3.1	2.4	2.7
σ_ψ	(deg)	3.6	4.9	6.2	6.6	3.6	5.5
RMS h	(ft)	12.5(15.4)	25.4	41.3	39.3	35.6	35.9
RMS u	(kts)	0.8(1.4)	5.2	6.5	7.3	7.0	6.6
RMS y	(ft)	58.7(43.2)	43.0	84.5	66.3	76.8	69.4
RMS $\dot{y}$	(kts)	4.1	3.7	4.7	4.6	4.0	4.3
σ_{δ_e}	(deg)	0.7	0.8	0.7	0.6	0.5	0.7
σ_{CP}	(deg)	0.9	1.2	1.2	0.5	0.3	0.9
σ_{δ_a}	(deg)	1.5	0.4	0.4	0.4	0.3	0.4
σ_{δ_r}	(deg)	1.2	0.9	1.0	0.6	0.4	0.8
Overall performance, J_m		0.05(0.07)	0.11	0.35	0.26	0.27	0.25
Replications		.	9	7	5	3	24

ALLOCATION OF ATTENTION	
θ	.19
$\dot{\theta}$	.21
φ	.08
$\dot{\varphi}$	.16
ψ	.07
$\dot{\psi}$	.11
h	.03
$\dot{h}$	.04
u	.03
y	.01
$\dot{y}$	.07
f_{LONG}	.5
P_O	-18 dB

(.): predictions with thresholds

Table 4: Model predictions and experimental results for the NAV-H task

PARAMETER		SUBJECT A MODEL	SUBJECT A MEASURED	SUBJECT B MODEL	SUBJECT B MEASURED
σ_θ	(deg)	2.1	1.7	2.1	1.6
σ_φ	(deg)	1.4	1.6	1.5	1.3
RMS ψ	(deg)	6.8	2.5	7.8	4.8
RMS h	(ft)	22.3	21.3	54.5	55.1
RMS u	(kts)	1.5	1.5	2.1	2.5
RMS v	(kts)	1.2	1.2	1.7	1.6
RMS v_h	(kts)	1.9	1.9	2.7	3.0
σ_{δ_e}	(deg)	1.0	1.1	1.0	1.0
σ_{CP}	(deg)	0.8	1.1	0.8	0.8
σ_{δ_a}	(deg)	0.2	0.6	0.14	0.5
σ_{δ_r}	(deg)	1.6	1.1	1.7	0.6
Overall performance, J_m		0.20	0.20	0.58	0.66
f_{LONG}		0.57		0.57	
P_O		-16 dB		-15 dB	
Threshold ratio TH		1/6 display limit		1/2 display limit	

Table 5: Model "match" and experimental results of the HOVER task

PARAMETER		MODEL "MATCH"	MEASURED	
			NAV-H	NAV-L
σ_θ	(deg)	1.9	2.6	2.1
σ_φ	(deg)	2.9	2.8	2.8
σ_ϕ	(deg)	3.6	4.9	4.7
RMS h	(ft)	24.0	25.4	23.9
RMS u	(kts)	2.5	5.2	5.1
RMS y	(ft)	44.6	43.0	46.6
RMS $\dot{y}$		4.1	3.7	3.9
σ_{δ_e}	(deg)	0.6	0.8	0.8
σ_{CP}	(deg)	0.6	1.2	1.1
σ_{δ_a}	(deg)	1.3	0.4	0.4
σ_{δ_r}	(deg)	1.0	0.9	0.8
Overall performance, J_m		0.11	0.11	0.11

ALLOCATION OF ATTENTION	
θ	.12
$\dot{\theta}$	.07
φ	.08
$\dot{\varphi}$	.30
ϕ	.07
$\dot{\phi}$	.14
h	.04
$\dot{h}$	.02
u	.00
y	.02
$\dot{y}$	.14
f_{LONG}	.25
P_O	-16.4 dB
TH	1/6

Table 6: Model "match" and experimental results of the NAV tasks for subject A

PARAMETER		NAV-H		NAV-L	
		MODEL	MEAS.	MODEL	MEAS.
σ_θ	(deg)	2.1	1.9	1.9	1.6
σ_φ	(deg)	3.1	2.4	3.3	2.8
σ_ϕ	(deg)	4.1	6.2	4.4	5.8
RMS h	(ft)	41.7	41.3	28.3	27.2
RMS u	(kts)	2.8	6.5	2.5	5.7
RMS y	(ft)	80.0	84.5	85.0	88.7
RMS $\dot{y}$	(kts)	4.7	4.7	5.0	5.1
σ_{δ_e}	(deg)	0.7	0.7	0.6	0.6
σ_{CP}	(deg)	0.6	1.2	0.8	1.0
σ_{δ_a}	(deg)	1.3	0.4	1.3	0.4
σ_{δ_r}	(deg)	1.0	1.0	1.0	0.8
Overall performance, J_m		0.33	0.35	0.26	0.27

ALLOCATION OF ATTENTION		
PAR.	NAV-H	NAV-L
θ	.12	.16
$\dot{\theta}$	.07	.09
φ	.08	.07
$\dot{\varphi}$	.30	.27
ϕ	.07	.06
$\dot{\phi}$	.14	.12
h	.04	.05
$\dot{h}$	.02	.03
u	.00	.00
y	.02	.02
$\dot{y}$	.14	.13
f_{LONG}	.25	.33
P_O	-15.4 dB	
TH	1/2	

Table 7: Model "match" and experimental results of the NAV tasks for subject B

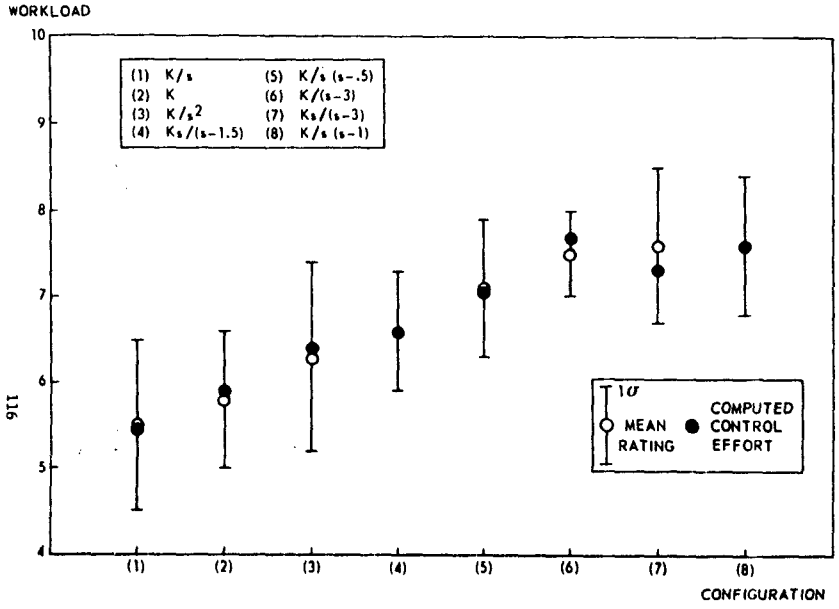
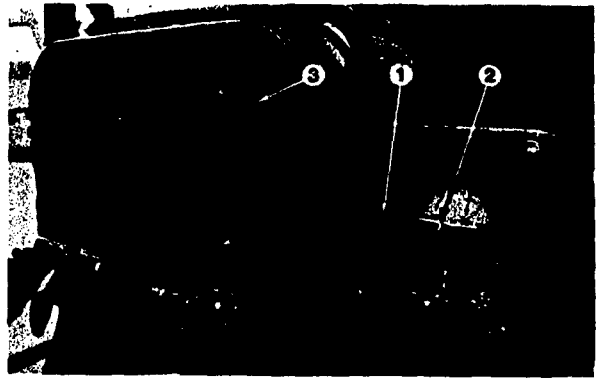


FIG. 2 A COMPARISON OF COMPUTED CONTROL EFFORT AND SUBJECTIVE RATINGS.(DERIVED FROM REF. 7)



1 HEIGHT ERROR
2 HORIZONTAL VELOCITIES (HOVER)
LATERAL TRACK DEVIATION (NAV)
3 INDICATED AIRSPEED (NAV)

FIG. 3 ALOUETTE III FLIGHT INSTRUMENTS

FIG. 4 ALOUETTE III HELICOPTER

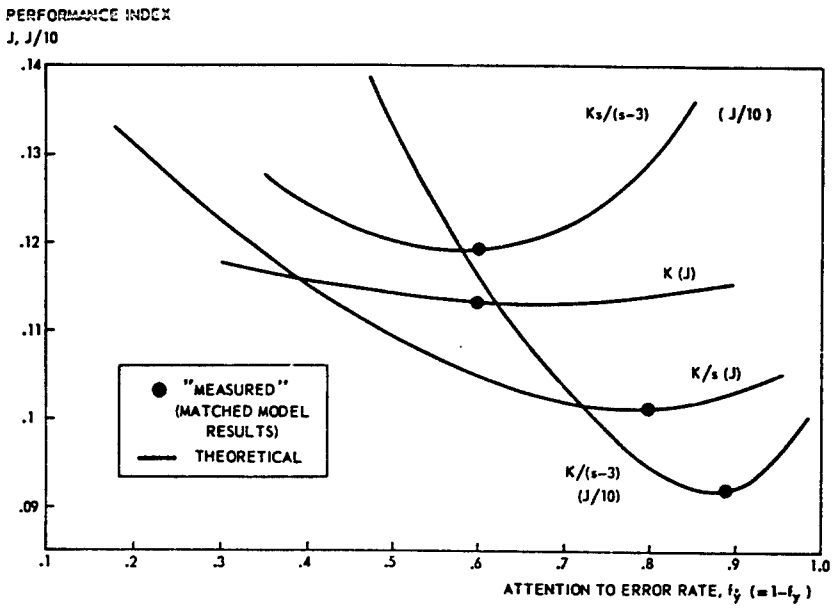


FIG. 1 ALLOCATION OF ATTENTION VERSUS PERFORMANCE-THEORETICAL AND MEASURED RESULTS. (DERIVED FROM REF. 7)

SUBJECT	TASK	COMPUTED EFFORT	SUBJECTIVE RATING		
			AVERAGE	EFFORT	DEMAND
A	HOVER	17.0	5.3	5.0	5.5
	NAV-L	15.9	4.9	4.5	5.4
	NAV-H	15.9	4.6	4.3	4.8
B	HOVER	16.5	7.5	7.5	7.4
	NAV-L	15.1	6.4	6.4	6.4
	NAV-H	14.9	6.7	6.7	6.7
Average of 4 subjects	NAV-L	15.0	5.3	5.2	5.5
	NAV-H	15.1	5.4	5.2	5.7

Table 8: Comparison of control effort model results and subjective ratings

FIG. 5 PERFORMANCE VERSUS CONTROL EFFORT FOR SUBJECT A.

MULTI-ATTRIBUTE SUBJECTIVE EVALUATIONS OF MANUAL TRACKING TASKS VS. OBJECTIVE PERFORMANCE OF THE HUMAN OPERATOR

by Alex Siapkaras
Conter for Space Research
Massachusetts Institute of Technology

ABSTRACT

Cooper-Harper ratings have been common measures for evaluating the response of aircrafts from the pilots' point of view. It has been noted, however, that in general the rendering of such subjective evaluations for the handling qualities of engineering systems can neither account for the multi-dimensional nature of tracking and/or monitoring tasks nor is it invariant to operator-centered variables, including the operator's perception of the task.

This paper develops a computational method to deal with these problems by defining matrix -rather than one-dimensional- ratings based on multi-dimensional scaling techniques and multi-variate analysis (no utility functions involved). The method consists of two distinct analytical steps: 1) to determine the mathematical space of subjective judgements of a certain individual (or group of evaluators) for a given set of tasks and experimental conditionings, and 2) to relate this space with respect to both the task variables and the objective performance criteria used. In this space there exist vectors along which the one-dimensional C-H normative scale (in McDonnell's sense of the term) or any multi-attribute utility function can be validly assessed.

Results from this method have been obtained for a variety of second-order trackings with smoothed noise-driven inputs, and they clearly show how: 1) many of the internally perceived task variables form a non-orthogonal set, and 2) the structure of the subjective space varies among groups of individuals according to the degree of familiarity they have with such tasks.

DEFINITION OF THE PROBLEM

In manual tracking tasks one deals with a man/machine system according to Fig.1. For a given human operator and extra-systemic conditions, the task is a blend of the requested mission, the display configurational mode and the dynamics of the machine or vehicle with its associated controlled elements.

$$task(\underline{x}, D, C) \qquad (1)$$

While the vehicle dynamics and type of controls can be easily characterized by certain parameters (like gain, damping ratio, etc), there are not even known and widely applicable parameters of statistical nature that are capable of characterizing feedback displays. Scalings, luminosity, and measures of legibility and information contents have been the focus of various people's efforts. For reasons of this basic lack of understanding, as well as for simplicity in the design of control/display configurations, it is customary to fix all pre-mentioned papameters to a certain level. On the other hand one requires maximum flexibility in what the input might be, so that the mission can be both time-varying and performance-dependent

$$\underline{X}(t, \underline{Y})$$

For a vehicle, a possible state description might be ($\underline{r}$ $\dot{\underline{r}}$ E), where $\underline{r}$ the position and E($\ddot{\underline{r}}$,ϵ,V) a measure of energy expenditure, environmental impact, effort allocation, and vehicle economics and kinematics. An associated constraint on permitted accelerations might be of the form $\ddot{r} < f(g,t,task)$; and so on, ... the model proceeding to its natural or unnatural conclusion.

In order to stabilize the performance of the system one feeds $\underline{\epsilon}$ back to the input

$$\frac{\underline{Y}}{\underline{X}} = \frac{1}{1 + G_b G_H G_c} \quad or \quad \underline{Y} = (1 + CHD)^{-1} CHD\, \underline{X}$$

so that essentially from the operator's point of view the mission becomes a function of the error and its past history

$$h_1(\underline{\epsilon}, D, \int(\underline{\epsilon})'s)$$

while his command actions are still governed [1] by

$$h_2(h_1, \phi, \psi, \eta, G_H).$$

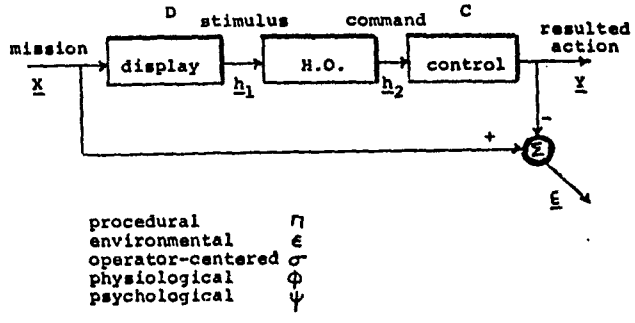


Figure 1

Given a set of N disjoint tasks, the operator can rate these manual tracking tasks in an internal frame of reference of his by comparing the "innermost experience" he encountered during their execution. The hypothesis done is that this comparison is not exclusively based on his objective performance, but is a verbalization of his inner state of mind, physique, mood, or whatever one might be inclined to call it, hopefully relevant to the execution of the tasks [2].

These comparisons have to be done in the most elemental of fashions, ie. by pairs, so that one can investigate the most minute details of preference, confusion, inconsistency, etc. characterizing the subjective evaluations rendered by the operator. Otherwise, a direct comparison of all N tasks will only reveal judgements on a one-dimensional basis [3]. How can one get into the internal structure of the operator's evaluation mechanism might be a metaphysical question, if it were not for gathering and interpreting these comparisons in a very engineering and mathematical sense respectively.

Comparison information is gathered in the form of a certain numerical measure rendered by the operator on a physical variable on which humans are known to exhibit a fair aptitude for precise estimates (for ex. length, frequency; not weight or time etc). In general, this judgement-to-numbers mapping performed by the operator is subject to the sequence of task pairs already presented and to the sensory and motor mechanisms employed, as dictated from the physical analogy used. Focusing on the psychophysical, rather than on the physiological aspects of this mapping, if the numerical value for the similarity between the tasks of the n^{th} pair is $d_{i_n j_n}$, then

$$d_{i_n j_n} = d_{i_n j_n}\left(\eta, C_{\epsilon, n-1}\right) \qquad (2)$$

Comparison information is interpreted as a measure of dissimilarity between the tasks of a given pair, each task being represented by a point in an abstract "subjective" space, the distance of the point from the rest being prescribed by the corresponding d_{ij}'s. The closer two points the more similar the tasks corresponding to those points are. The multi-dimensional nature of the manual tracking task suggests that in order to fit all those points in their appropriate position within the space one has to resort to a space of enormous dimensionality. If there are N different tasks compared to each other, then in general R^{N-1} would be a sufficient space to accommodate without any discrepencies all points of interest subject to their mutual constraints, as dictated by their distances.

The particular form of the metric used to denote distances in this space, of the coordinate representation used for labelling the points in it, and of the reduction used to alleviate unnecessary dimensions, is an altogether different matter. (See discussion under "The Subjective Space").

THE EXPERIMENTAL PAPAMETERS

According to definition (1) each task is characterizable from three entities. For simple one-dimensional compensatory tracking tasks displayed on a CRT in standard proportional fashion, these reduce to two: the mission and the controls. Therefore, there are two kinds of parameters needed to specify a task: standard control parameters and parameters capable of describing the essential elements of the mission.

Fig 2 shows the task space comprised out of the well-accepted parameters for second order plants

$$G_c = \frac{K}{(s/\omega_n)^2 + 2(\zeta/\omega_n)s + 1}$$

In contrast to the natural frequencies and the damping ratios, the gains are not prespecifiable; they are rather derived as a consolidation between input strength and display dimensions.

$$\left[\int_0^T\left[\int_0^t g_c(t')dt'\right]^2 dt\right]^{1/2} = T^{1/2}(Z)L \Big/ S_\epsilon^{1/2}(T;0) \qquad (3)$$

which says that the RMS output due to an RMS input be the same specified level for all tasks (where T a certain duration of the unobstacted signal and L the height of the screen). Of course, if the LHS is an unbounded quantity, a meaningful selection of T is of critical importance.

Choosing parameters to represent the mission is a more difficult task, even for one-dimensional trackings. It involves two aspects: specifying a nominal path and superimposing a certain disturbance. Because many of the recent experiments involve "non-sensical" rather than smooth slow-varying deterministic inputs, this study was compelled towards representing the input signal in terms of a statistical description which has also the property of becoming an exact representation for the case the inputs are non-stochastic.

Starting from a Taylor-like expansion for the property f associated with the N^{th} value of the running process x(t)

$$f(x_N) = Nf(r) - \sum_{\kappa=1}^{N-1} f(x_\kappa) + N\sum_{k=1}^{\infty}\frac{d^k f(r)}{dx^k}\,\frac{S_k(N;r)}{k!}$$

one can express the Fourier series of a continuous process n(t) in terms of the statistics S_k and a reference time t_r, by

$$\eta(t) = \frac{1}{2}\sum_{\kappa=-\infty}^{\infty}\left\{\eta(t_r) + \sum_{k=1}^{\infty}\frac{C_{\kappa k}(t_r)}{k!}S_k\left(t;\eta(t_r)\right)\right\} e^{i\omega_\kappa(t-t_r)} \qquad (4)$$

where $\omega_k = \frac{kn}{t}$, and C_{kk} a recursive coefficient with $C_{k0} = \eta(t)$ and

$$C_{\kappa k} = 1 - i\omega_k \frac{C_{k,k-1}}{\dot{C}_{k,k-1}}.$$

This implies that even if one knew the exact time-history of the statistics one would not in principle reproduce exactly the input signal, since for infinitely differentiable $n(t)$'s one would truncate the series. However, the power of the above result is that, if one could start specifying some of the properties on frequency rather than just on amplitude, one could conceivably acquire information about the signal in a faster rate. For example, there have been in the past pseudo-random trackable signals generated only by combining a few sinuisoids with random mixing of amplitudes. No series on amplitude statistics involved, just a randomly selected $n(t_r)$ which is different for each ω_k.

Returning to our general result of (4), the question arises: which of the ω_k's are most essential, beyond specifying certain of the S_k's ? In other words, what is the minimum prescription for defining a frequency in a signal which has no apparent primary mode of oscillation? It is suggested in this paper that such a primitive notion of frequency is provided by the average rate the signal crosses its mean. The RMS_ω of fig 2 signifies such a frequency.

It can be noted that this primitive frequency is a function of time both explicitely and implicitely:

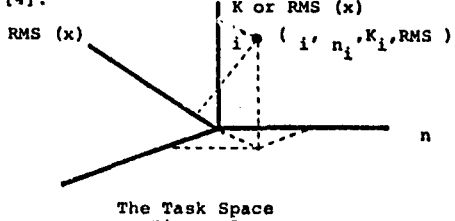
$$RMS_\omega(t, \varepsilon, (t))$$

Explicitely, because the rate of crossings depends from the incoming signal flow; implicitely, because the mean of the signal itself changes. For a purely random sequence one would of course expect the mean to be constant, and the incoming signal to have an equal chance to go over or under the mean, in which case the frequency remains on the average constant.

Summarizing the above, one could say that all non-procedural experimental parameters are incorporated in the notion of task variables which, in the case of one-dimensional tracking on fixed display format and proportional imaging, are:

$$x = \{S_k(t; n(t_r)) \text{ with } k = 1, \dots \}$$
$$D = \{K(L) \text{ of } G_c \text{ as given by (3)}\}$$
$$C = \{a_k\text{'s of } G_c = a_{-m}s^{-m} + \dots + a_0 + \dots + a_n s^n\}$$

or combinations thereof. For any practical reasons $k_{max} = 2$ and $m_{max} = n_{max} = 2$ for PID plants. An interesting application where there is need to keep more of the statistics S_k characterizing x is treated in [4].



The Task Space
Figure 2

After having defined the nature of the non-procedural experimental parameters, comes the time to talk about the procedural parameters of the experiment. Definition (2) explains the object of the experimental sessions: the rendering of subjective judgements about the similarity of the tasks of given pairs in a numerically quantifiable form. When is the sequence of pairs more inconsequential to both the objective performance and the subjective evaluations of the operator? A standard sequence of randomly mixed pairs with each pair appearing at least two times seems to be a satisfactory answer.

However, this approach inevitably leads to long sessions where fatigue and boring play a very important role, and induces the tendency for insufficient in-between-tasks time of de-adaptation and for ridiculously low task durations. For example a 2-hour session with three values for each task variable of fig.2 would result in a task duration of ... .55 secs! Besides, in such an environment of abstract simulation and of intense need for continous re-adaptation, the subjective sense of time for the human operator is severely offset [5].

A heuristic selection from the set of possible combinations of task variable values seems a much sounder method than an arbitrary splitting up of the experimental session. The selection can be achieved via the idea of derangements and partitions [6]. Of course, this way we get far less d_{ij}'s than necessary to fix our R^{N-1} space, so we need to extract d_{ij} from the human operator other kinds of comparison information as well. For example, a measure of combined difficulty, or any other characteristic f_{ij}, in addition to similaritios [4]. This type of experimental procedure would constitute a "fractional replication factorial design" according to [7].

THE SUBJECTIVE SPACE

Now, it is time to go back to the concluding remark made on the definition of the problem. How is the numerical information, gathered from the operator comparing manual tracking tasks in pairs, mathematically interpreted as a function $f(x_1(i), \dots x_{N-1}(i); x_1(j), \dots x_{N-1}(j))$ of certain "hypothesized" N-1 coordinates x_{N-1} in the abstract space of subjective evaluations or "subjective space"? The essence of multi-dimensional scaling techniques lies in interpreting dissimilarity as a distance d_{ij}, and distance as a Minkowski metric [8]:

$$d_{ij} = \left[\sum_{k=1}^{N-1} |x_k(i) - x_k(j)|^p \right]^{1/p}$$

For Euclidean spaces p=2 and in order for the measured d_{ij}'s to be true physical distances one requires that d_{ij} is never greater than $d_{ik} + d_{kj}$. However there is no guarantee that this will be the case. The standard transformation used to ensure that this will hold is given by [9]:

$$\delta_{ij} = d_{ij} + \text{OFFSET}$$
$$\text{where OFFSET} = \max_{\{i,j,k\}} (d_{ij} - d_{ik} - d_{kj} > 0)$$

Since we know $\mathcal{N}$ dissimilarity measurements δ_{ij} and at the same time have [N tasks * (N-1) coordinates] = $2\mathcal{N}$ unknowns, there are $\mathcal{N}$ degrees of freedom. [4] specifies an other $\mathcal{N}$ constraints by taking into account symmetric measurements f_{ij} on an extra comparison measure. [10] adds constraints that fix the origin at the centroid of all N points, and point the axes at certain orientations. The advantage over [4] and [10] of the method presented below is that one can obtain directly a closed form solution when setting the upper triangular table of fig. 3 to zero: the first point is at the origin, the second is on one of the axes, the third on any of the planes that incoprorate the previously considered axis, and so on.

$$x_k(i) = \frac{\delta_{ki}^2 - \delta_{(k+1)i}^2 + x_k^2(k+1) - \sum_{r=1}^{k-1} \left\{ \left[x_r(k+1) - x_r(i) \right]^2 - \left[x_r(k) - x_r(i) \right]^2 \right\}}{2\, x_k(k+1)} \qquad (5)$$

Equation (5) is the coordinate representation derived for labeling the point i, and is based on a Euclidean metric. It computes the coordinate values in terms of the coordinates $x_r(k)$ of previously computed points, and of previously computed coordinates $x_r(i)$ of the point in question.

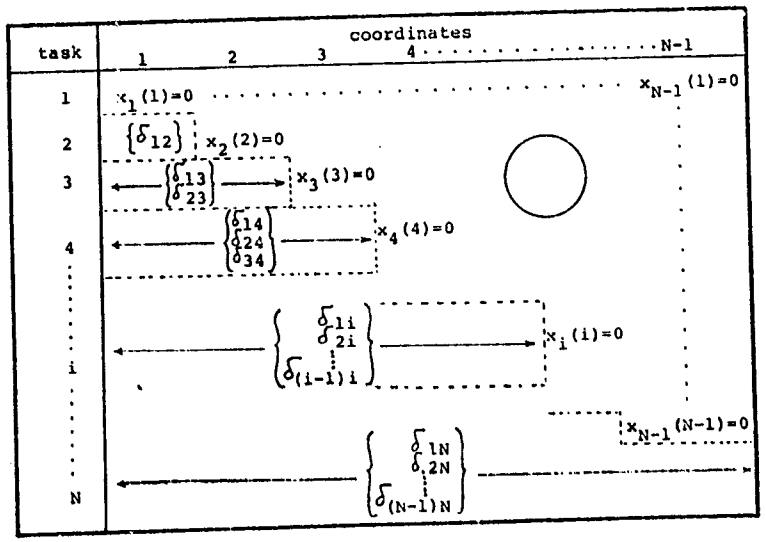


Figure 3

Since the dimensionality of the subjective space far exceeds the total number of task variables, one would like to arrive at a reduced space representation of the essential character portrayed by the"exact space" with much fewer dimensions. [11] lists a number of methods that do one of three things: either trim the exact space from low RMS coordinates and compensate the remaining of it with a best fit rearrangement, or partition the space in regions of high concentration and derive a lower order space corresponding to only the distances between the centroids of these clusters, or directly apply a statistical hyperplane fit(of prespecified dimensionality) to the exact space. The advantage over the methods listed in [11] of the method presented below is that the procedure outlined is exact and non-iterative in nature.

First one reorders the coordinates in order of least importance $\{x's\} \rightarrow \{\tilde{x}'s\}$. Looking at figure 3 and accepting that the significance of coordinate k is the probability that a point has a non-zero k^{th} coordinate times the average relative importance of coordinate k for all points i that have values on k:

$$\text{SIGNIFICANCE OF COORD k} = \frac{N-k}{(N-1)(N-1)/2+N} \cdot \frac{\sum_{i=k+1,N} \frac{x_k^2(i)}{E_2(i-1)i}}{N-k} \propto \sum_{i=k+1}^{N} \frac{(i-1)\,x_k^2(i)}{\sum_{l=1,i-1} x_l^2(i)}$$

Secondly, one would like to define transformations on $\tilde{x}'s$ such that R linear combinations of them ξ are equal to zero. This then would mean that all information contained in $\mathcal{R}^{N-1-R}$ will be contained after the transformations are found in $\mathcal{R}^{N-1-R}$. Suppose $f_k(\vec{x}_k)$ the tranformation rule on the k^{th} coordinate and T the matrix that maps the basis $\{\underline{e}_{N-1}\}$ of $\mathcal{R}^{N-1}$ to a new basis $\{\underline{e}'_{N-1}\}$ in a space $\mathcal{R}^{N-1}$ which contains the sought after $\mathcal{R}^{N-1-R}$. Then our complete description of the mapped transformation is:

$$\begin{pmatrix} \xi_1\, \underline{e}'_1 \\ \xi_2\, \underline{e}'_2 \\ \vdots \\ \xi_{N-1-R}\, \underline{e}'_{N-1-R} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \} \mathcal{R} = T \begin{pmatrix} f_1(\tilde{x}_1)\, \underline{e}_1 \\ f_2(\tilde{x}_2)\, \underline{e}_2 \\ \vdots \\ f_{N-1}(\tilde{x}_{N-1})\, \underline{e}_{N-1} \end{pmatrix} \uparrow \text{order of greater importance}$$

Thirdly, one would like if possible to have an exact description on the new coordinates $\xi_1, \xi_2 \cdots \xi_{N-1-R}$ in terms only of the transformations f of the most important coordinates, in other words of $f_i(\tilde{x}_i)$ rather than of $f_{i+1}(\tilde{x}_{i+1})$. This implies that T should be lower triangular-like in form. Fig 4 indicates how such a T can be formed. Each row k represents a unit vector $\underline{e}'_k$ which should be orthogonal to all others, ie. $\underline{e}'_k \cdot \underline{e}_k = \delta_{kk}$.

$$\tau_{ii} = -\left(\alpha^2 + \sum_{k=2}^{i-1} \tau_{kk}^2 \right) \Big/ \tau_{i-1,i}$$

$$\tau_{i,i+1} = \left[1 - \alpha^2 - \sum_{i=2}^{i} \left| \tau_{kk} \right|^2 \right]^{1/2}$$

and $\left\{ \begin{array}{l} \varepsilon\left[\alpha^2 + \sum_{K=2}^{N-2} \tau_{KK}^2\right] + \beta\tau_{(N-1)(N-1)} = 0 \\ \varepsilon^2\left[\alpha^2 + \sum_{K=2}^{N-2} |\tau_{KK}|^2\right] + \beta^2 = 1 \end{array} \right\}$ determine ε and β.

Fourthly, one would like to have meaningful transformations f's. Polynomial stretching-shrinking functions $\tilde{x}_K = \sum_{\ell=0,Q} a_{K\ell}\, \tilde{x}_K^{\ell} + x_K^{Q+1}$ that preserve the order of the points in the untransformed space are used. Suppose O_{Ki} the order of the projection of the i^{th} point on the k^{th} coordinate, then

$$O_{Ki} - \tilde{x}_K^{Q+1}(i) = \sum_{\ell=0}^{Q} a_{K\ell}\, \tilde{x}_K^{\ell}(i)$$

are N(N-1) equations for the (N-1)(Q+1) unknown coefficients $a_{K\ell}$. The rest RN equations come from requiring that $\xi_K = \sum_{\sigma=1}^{N-1} \tau_{K\sigma} f_\sigma(\tilde{x}_\sigma) = 0$ for all new coordinates beyond the desired reduction; in other words

$$-\sum_{\sigma=1}^{N-1} \tau_{K\sigma}\, \tilde{x}_\sigma(i) = \sum_{\sigma=1}^{N-1} \tau_{K\sigma}\, a_{\sigma\ell}\, \tilde{x}_\sigma^{\ell}(i) \quad \text{for } K = N-R \text{ to } N-1.$$

For a unique determination of the coefficients one has to put $Q = N + R + 1$ and specify arbitrarily a total of $2(N-1)-R$ of them. After all a's have been determined, one arrives at the reduced

$$T = \begin{bmatrix} \alpha & \sqrt{1-\alpha^2} & & & & \\ \alpha & \tau_{22} & \tau_{23} & & & \\ \alpha & \tau_{22} & \tau_{33} & \tau_{34} & & \\ \vdots & & & \ddots & & \\ \alpha & \tau_{22} & \tau_{33} & \tau_{44} & \cdots & \tau_{(N-2)(N-1)} \\ \varepsilon\alpha & \varepsilon\tau_{22} & \varepsilon\tau_{33} & \varepsilon\tau_{44} & \cdots & \varepsilon\tau_{(N-2)(N-1)} \quad \beta \end{bmatrix}$$

Figure 4

subjective space

$$\xi_K = \left\{ \sum_{\sigma=1}^{K+1} \left| \tau_{K\sigma} \cdot \left(\sum_{\ell=0}^{Q} a_{\sigma\ell}\, \tilde{x}_\sigma^{\ell} + \tilde{x}_\sigma^{Q+1} \right) \right|^2 \right\}^{1/2} \qquad (6)$$

which is an exact representation in terms of the original subjective space. Information from the least significant of the coordinates $\tilde{x}$ is not lost, because that information is carried implicitely in the determination of the coefficients.

PROPERTY VECTORS

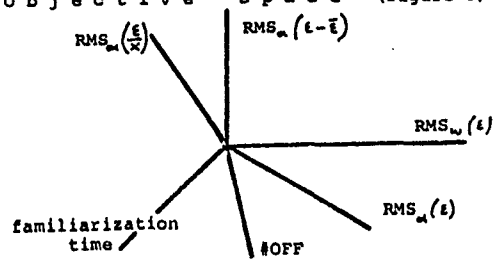
Having arrived at a reduced space for subjective evaluations, with a dimensionality $N - 1 - R = M$ comparable to the hypothesized independent task variables, one would like to know on which directions in this space a certain property P of the tasks experiences maximum change.

$$\left[\nabla P(\xi_1, \xi_2 \cdots \xi_M)\right]_{max} \text{ implies } \sum_{K=1}^{M} \frac{\partial P}{\partial \xi_K} \frac{\partial^2 P}{\partial \xi_j \partial \xi_K} = 0 \quad \forall\, j = 2, M$$

where ∇ denotes the gradient in the reduced subjective space. There are three issues involved here: which are the properties P of interest; how does one arrive at their functional form in the reduced space formulation; and which are the best linear functions $\xi_M(\xi_1), \xi_{M-1}(\xi_1), \ldots \xi_2(\xi_1)$ that represent the vector along which such a change occurs.

There are only three kinds of task properties one could think of:

(a) the task variables (see discussion under "The Experimental Parameters")
(b) various functions of subjective nature. The normative Cooper-Harper scale [3], multi-attibute utility functions [11], etc.
(c) various measures of the operator's performance. These are observed and computed statistics characterizing the objective performance of the human operator on any given task, and they comprise the O b j e c t i v e S p a c e (figure 5).



The Objective Space
Figure 5

There are two classes of interesting statistics $S_k(t;r)$ that characterize the objective performance, statistics on $\epsilon(t)$ and statistics on $\epsilon(t)/x(t)$. Suppose $\phi(\epsilon(t),x(t))$, then in order for one to know any

$$\frac{1}{t}\int_0^t \phi^k(\tau)d\tau = E_1^k(t) + \sum_{k=1}^{K} \frac{k!/k!}{(k-k-1)!} E_1^{k-k}(t)\, S_k\left(t;E_1(t)\right)$$

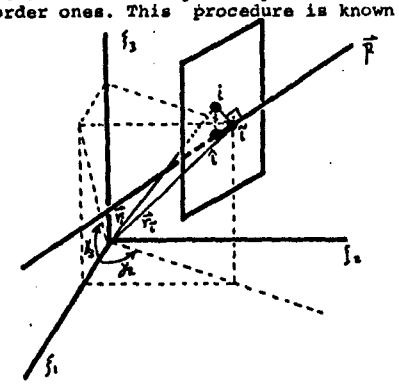
one also needs $E_1(t)$. Figure 5 portrays the case k=2, along with the RMS frequency at which the error becomes + or -.

In addition to the measures that are, one way or another, statistical functions of the error and the input h_1 (in our case $\underline{x}$), there are other performance measures which are display-dependent or non-accumulative in character. An example of the first kind is the number #OFF of times the display output gets off the limits of the CRT screen. An example of the second kind is the time required for the human operator to remain within a certain percentage of failure from exceeding the standard deviation achieved by task's end.

Having defined above what is meant by property in the context of psychophysical engineering, one can address the second question: how does one arrive at their functional form. Given, for example, C-H ratings corresponding to the N tasks, one has N equations to solve for a total of N suitably selected a's.

$$P(i) = \sum_{\{K\}} \alpha(K_1 K_2 \ldots K_M)\, \xi_1^{K_1} \xi_2^{K_2} \cdots \xi_M^{K_M}$$

The sequence in which one would employ the a's should be governed by the rationale that lower order Taylor-expansion-like terms are prefered to higher order ones. This procedure is known as multivariate synthesis.



Property Fitting
Figure 6

Now suppose $\tilde{i}$ is the projection of the point i, corresponding to the i^{th} task in the reduced subjective space, on the sought after vector $\vec{P} = \xi_1' + \sum_{k=2}^{M} \gamma_k \xi_k'$ on which $\xi_k = \bar{\xi}_k + \gamma_k + \delta_k$ and which minimizes the cost functional $J^2 = \sum_{i=1}^{n} [P(\tilde{i}) - P(i)]^2$. By combining the relation $(\vec{\tilde{r}}_i - \vec{r}_i)\cdot\vec{P} = 0$ (which renders $\tilde{i}$ the projection of i) with the idea of neglecting higher order terms in the expansion

$$P(\tilde{i}) = P(i) + \sum_{j=1}^{M} \frac{\partial P}{\partial \xi_j}\Big|_i \left(\bar{\xi}_j(i) - \xi_j(i)\right) + \sum_{k=2}^{\infty} \frac{1}{k!}\left[\sum_{j=1}^{M} (\bar{\xi}_j - \xi_j)\frac{\partial}{\partial \xi_j}\right]^k P$$

one arrives at expressions for $\partial J^2/\partial \gamma_k = \partial J^2/\partial \delta_k = 0$ that can be solved iteratively to find the γ's and δ's which best specify the direction (the γ's) and the origin (the δ's) of $\vec{P}$.

The advantage of this method over the methods mentioned in [11] is that one fixes completely the property vector in the reduced subjective evaluation space, while the rest assume that in all regions of the space the property direction keeps always parallel to itself. Also, as one can inspect from figure 6, the best estimates $\tilde{i}$'s for the points are exact projections on the fitted vector rather than mere images $\hat{i}$'s on planes parallel to the f_M axis. If $\vec{P}_1, \vec{P}_2, \ldots \vec{P}_L$ the property vectors corresponding to the task variables, then one can derive the matrix $\prod_{L\times M}$ that defines the mapping from the subjective space to the task vectors, and call $\prod$ the "matrix rating" associated with this subjective space.

DISCUSSION

It is this last technique that allows us to find out if the mapping of the task space in the subjective space maintains the orthogonality between the task vectors or not. If, for example, we arrive at an experimentally deduced law of universal validity about certain task variables not comprising an orthogonal set in the subjective space, then it might mean only three things:

(a). In our abstract modelling of the task we might have misconceived or altogether skipped certain interdependencies amongst the so-called task variables.

(b). In our internal model of the task we might have misperceived, -or inadequately learned to appreciate-, the independence of the task variables.

(c). Mathematical misformulations and psychophysical misjustifications concerning developments in the theory of subjective spaces.

I would like to suggest that all three factors come into play:

(A). Not accounting for the higher order statistics in the representation of the input disturbance, in the new parameters introduced by the digitization of the control structures in the simulation environment, etc.

(B). The human operator might subconsciously introduce artificial dependencies among the various independent parameters in order to facilitate data reduction and information storage mechanisms that might govern his/her internal response and decision processing.

(C). Dissimilarity judgements might not be modellable by a triangular inequality, so that their mathematical interpretation as distances might not be appropriate. But even if that is not the case, some Minkowski metric might exist such that it does preserve the orthogonality of the mapped task space.

Independently of the above fundamental questions, however, there is much utility to be found in exploring the subjective evaluation space of human operators in manual tracking tasks.

(α). Knowing a sufficient description of the reduced subjective space corresponding to such and such manual tracking tasks, one could predict the subjective evaluations and objective performance of any human operator whose psychophysical aspects are similar to those for which the particular space was constucted.

(β). One could find for which categories of people such and such task variables are not treated independent of each other.

(γ). The subjective space could become a ground for testing the objective performance indices that best co-allign with the C-H scales.

Actually, this last issue is one of the most interesting. It raises the possibility of resolving the often observed incompatibility between objective performance and subjective evaluation. What one would expect from a rational processor like the human operator is to match his internal satisfaction with the task to his actual performance. There is speculation, however, about three possible antagonistic mechanisms present that could be checked via the subjective space idea:

(A). Improving one's own satisfaction with a particular task does not necessarily mean better objective performance.

(B). Increasing the disheartening aspects of a task might actually result in improved performance.

(Γ). Teaching the human operator to consider the right type of variables for his/her task -by utilizing the mapping of the task space in his subjective space- might result in a worse optimal control representation for his internal model, due to (B).

If such antagonisms between subjective evaluations and objective performance are indeed sometimes observed, one should rest easy; according to [12] these antagonisms often result from acceptable "rational" kinds of task conceptualization and internal goal-defining.

NOMENCLATURE

c_{kn}	the sequence of the first n k-tuples presented $\{(i_1,j_1,\underbrace{k_1,\ldots}_{k}),(i_2,j_2,k_2,\ldots),\ldots.(i_n,j_n,k_n,\ldots)$
(i,j)	the pair comprised of the i^{th} and j^{th} tasks
N	the total number of tasks $\prod_{I=1}^{L} n_I$ (L=# of task variables)
$\mathcal{N}$	all possible pairs from N tasks N(N-1)/2
R^m	an m-dimensional space
n_I	the # of the different values on the I^{th} task var.
$E_k(N)$	the k^{th} average of the first N values, ie. $\overline{x^k}$ $E_k(N) = x_N^k / N + (N-1)\, E_k(N-1) / N$
$S_k(N;r)$	the statistic $\langle (x-r)^k \rangle_x$ of an N-member sampling x from r $S_k(N;r) = \sum_{k=0}^{K} \frac{(-1)^{K+k}\, K!\, E_r^{K-k}(r)}{k!\,(K-k)!}\, E_k(N)$
$\Theta_n(N)$	the n^{th} threshold of a regressional filter $\theta_n(N) = \theta_{n-1}(N) - \frac{\lvert s_n(N;\, n-1)\rvert^{1/n}}{2^{n-2}}$ with $\theta_1(N) = E_1(N)$
$T(\underline{e}_m;\underline{e}'_n)$	the tranformation matrix that maps the coordinate system of an original R^m space to a new R^n space whose coordinate system representation is in a basis $\underline{e}'$.

REFERENCES

[1] McRuer + Krendel "Mathematical Models of Human Pilot Behavior" NATO Agard-AG-188, Jan 1974

[2] A.Siapkara "The Cognitive Processes Underlying Human Pilot Models" 1975 (Unpublished. Can be obtained by request by writing to POB 6039,Boston,Ma 02209)

[3] J.McDonnell "Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities" AFFDL TR-68-76,Dec 68

[4] A.Siapkara "Modelling and Logic behind the Evaluation of Automated Production Centers" 1976 (Unpublished. Can be obtained by request from POB 6039,Boston,Ma 02209)

[5] J.Holubář "The Sense of Time" The State Medical Publishing House, Prague 1961

[6] S.Even "Algorithmic Combinatorics" Ed. Collier-Macmillan, 73

[7] D.Campbell + J.Stanley "Experimental and Quasi-experimental Designs for Research" The American Educational Research Association, 1963

[8] Kruskal,J.B. & F.J.Carmone "M-D␣SCAL␣5" Bell Telephone Laboratories,March 1969

[9] Torgerson,W.S. "Theory and Methods of Scaling" Wiley Ed., N.Y. 1960

[10] Carroll + Chang "INDSCAL analysis via an N-way Generalization of Eckart-Young Decomposition", Psychometrica 35, 1970, 283-319

[11] Paul E. Green + Yoram Wind "Multi-attribute Decisions in Marketing" The Dryden Press (chapters on multi-dimensional scaling)

[12] D.Bobrow + A,Collins "Representation & Understanding", Academic Press, 1975 (p 131-184).

THE EFFECTS OF PARTICIPATORY MODE AND TASK WORKLOAD ON THE DETECTION OF DYNAMIC SYSTEM FAILURES[1]

Christopher D. Wickens
Colin Kessel

Department of Psychology
University of Illinois at Urbana-Champaign

ABSTRACT

The ability of operators to detect step changes in the dynamics of control systems is investigated as a joint function of, (a) participatory mode: whether subjects are actively controlling those dynamics or are monitoring an autopilot controlling them, and (b) concurrent task workload. A theoretical analysis of detection in the two modes identifies factors that will favor detection in either mode. Three subjects detected system failures in either an autopilot or manual controlling mode, under single-task conditions and concurrently with a "subcritical" tracking task. Latency and accuracy of detection were assessed and related through a speed-accuracy tradeoff representation. It was concluded that failure detection performance was better during manual control than during autopilot control, and that the extent of this superiority was enhanced as dual-task load increased. Ensemble averaging and multiple regression techniques were then employed to investigate the cues utilized by the subjects in making their detection decisions.

INTRODUCTION

Over the past decade, the aviation industry has witnessed a gradual change in the role of the pilot in the cockpit. Many traditional pilot functions have been replaced by on-board computers, and in some instances the pilot is no more than a supervisor (Sheridan, 1976) or monitor of automatically controlled functions. One task, however, that remains of critical importance to the operator of any aviation system, whether he is removed from the control loop or not, is that of monitoring all facets of aircraft performance for the occurrence of failures or malfunctions. The relatively low frequency of occurrence of such events does not diminish the importance of failure monitoring and detection, because the consequences of an undetected malfunction, or one that is detected after an unnecessary delay, can be disastrous, potentially resulting in the loss of the aircraft or of human life.

[1]Contractual support for this research was provided by the Life Sciences Program, Air Force Office of Scientific Research, Contract Number F44620-76-C-0009. Dr. Alfred Fregly was the Scientific Monitor of the contract. The authors acknowledge the invaluable programming assistance of Mr. Roger Hunt in the conduct of this research.

It can be argued in fact that one criterion that should be used in considering whether a pilot should remain in the control loop under particular conditions is his relative sensitivity to system malfunctions in the two modes of participation.

Young (1969) has argued strongly on the basis of his findings that the operator is more sensitive to system malfunctions as an active participant in the control loop, than as a passive monitor. In his experiment, subjects were required to detect various step changes in system order and gain. Conditions were compared in which the subject was an active controller and a passive monitor (who was observing the compensatory display produced by another active controller). Under these circumstances detection latencies varied from twice to five times greater for the monitor than the controller. A second study which also compared detection ability in the two modes, however, resulted in contradictory findings. Ephrath (1975) investigated failure detection performance in a two-dimensional simulated landing task as a joint function of participatory mode and workload. The "failures," which in this case were deviations introduced into the flight path rather than changes in system dynamics, could occur in either the pitch or yaw channel. Under different conditions subjects were either in control when a failure occurred or were monitoring a nonadaptive autopilot in control of that channel. The non-failed channel could also be either controlled or monitored. Ephrath's results indicated a clear superiority for detection on the monitored as opposed to the controlled dimension, both in terms of the smaller number of missed failures and of the shorter detection latency. This difference Ephrath attributed in large part to the increased level of workload involved in the controlled task.

Obviously, in many respects the studies of Young and Ephrath are not comparable. Young employed single-axis tracking with changes in system dynamics, while Ephrath employed dual-axis simulator control with "deviation" failures. In addition, the monitoring conditions were different in the two experiments, being influenced by adaptation in Young's study and not in Ephrath's. In this light, it is not surprising that the conclusions differed dramatically. Certainly, one of the most salient differences between these studies lies in the contrast between single- and dual-axis tracking and is inherent in the greater workload imposed in the latter condition.

While numerous other investigations of failure detection performance are present in the literature (Sheridan and Johansen, 1976; and Young, 1969), the studies of Young and Ephrath are the only two that have explicitly contrasted detection between the two modes, so that a direct comparison is possible. The present study was conducted with the intent of clarifying the nature of the superiority relation between the two modes. A question of specific interest was whether the difference in results between the results of Ephrath's and Young's study could be attributable to differences in concurrent task workload between the paradigms, and for this reason secondary task workload was manipulated orthogonally to participatory mode.

Theoretical Analysis of Failure Detection

The detection of a failure or change in the characteristics of a dynamic system requires that the detector have available two basic elements: (1) an internal representation of the state of the normally operating system - the expected value of state variables and their expected variability, (Veldhuyzen and Stassen, 1976; Pew, 1974) and; (2) a channel, or set of channels, of information concerning the current state of the system. Failures are detected when the information concerning the current system state is assessed to be sufficiently deviant from the representation of normal operation to warrant a decision. The decision process involved may be assumed to involve the application of some statistical decision rule (Curry and Gai, 1976). The following theoretical analysis, employing the conceptual framework described above and represented symbolically in Figure 1, will attempt to define the characteristics or attributes of each participatory mode that might be expected to enhance the sensitivity of failure detection in that mode.

Stability of internal model of dynamics. An evolving conception in control theory is that the operator maintains an internal representation or 'model in the head" of the dynamic system that is being controlled. It is assumed here that this model provides the basis for predicting expected system outputs in response to known inputs - an internalized estimate of the transfer function of the system being controlled. This conception is consistent with that employed by Curry and Gai (1976), Miller and Elkind (1967), and others. With respect to failure detection, a critical characteristic of an internal model relates to its internal consistency or expected variability. For any given input to the system, the range or variability of possible outputs is a measure of this consistency.

It is proposed that when the operator is actively controlling, the stability of this internal model is considerably greater than when he is monitoring. This difference reflects the fact that, when controlling, the operator has a greater involvement with the system, and a direct knowledge of its input-output characteristics available by comparing his control inputs with the system response (Gould and Fu, 1966). This information is only available when monitoring if the monitored display is pursuit, and even then, knowledge of control inputs is not as precise, since a system response due to regulatory error correction, cannot easily be discriminated from one resulting from external disturbances. Thus provided with a smaller variance estimate of the normal state, detection in the control mode should show greater sensitivity to departures from this state induced by changing dynamics, than should detection in the monitoring mode.

Information channels. A second attribute of the control mode that predicts superiority of failure detection is the greater number of channels of information concerning the current state. When monitoring, information is provided to the operator exclusively via the visual channel (system error and its derivatives in the compensatory display or input and outputs

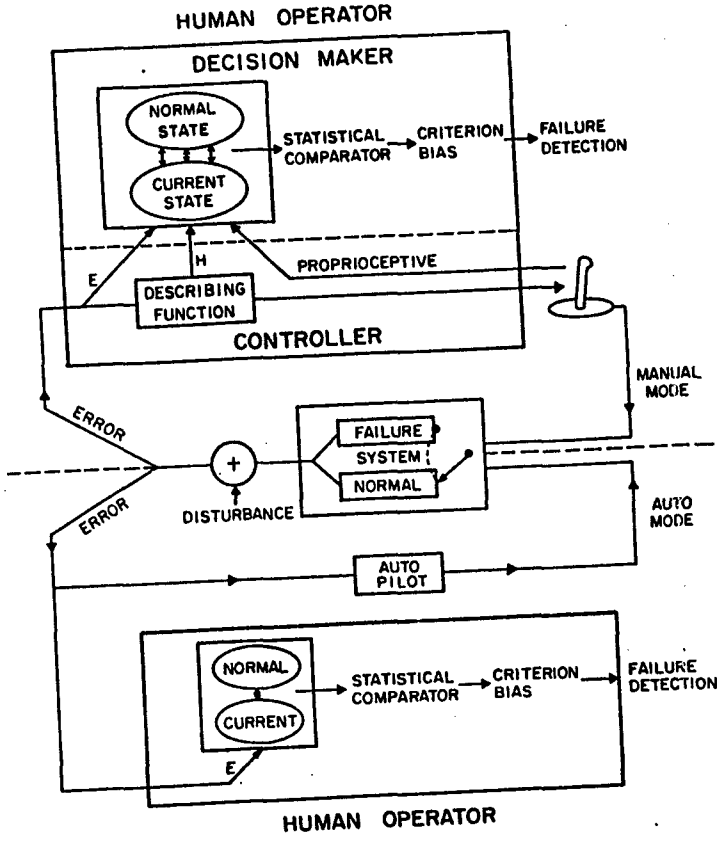


Figure 1. Schematic representation of human failure detection in manual mode (top) and autopilot mode (bottom).

plus derivatives in the pursuit display). On the other hand, in the control situation the operator also has available a proprioceptive channel of information concerning his own input to the control stick, independent of disturbances acting upon the system.

Although control input cannot directly reflect the occurrence of failures (except as failures initiate mechanical feedback from the control itself), it will do so indirectly to the extent that any compensatory adaptation that the operator initiates to a system change will be reflected in a change in his response characteristics (mean control position, velocity or acceleration) and/or the characteristics of the operator's open-loop transfer function. When controlling then, these proprioceptive channels will be available to the detection system to supplement the visual channels that are available in both monitoring and controlling (Figure 1).

While the controlling mode thus seemingly provides a distinct advantage over the monitoring mode by virtue of its added proprioceptive channel, it should be noted that this advantage is not invariably present for reasons related to the non-independence of control input and error. More specifically, if adaptation to the failure is rapid and complete, as may occur for example in response to shifts in system gain (Young, 1969), the obtained distribution of error following the change would show little or no alteration from that characterizing the normal operating state, while a change would be manifest in the characteristics of the control response and, therefore, the transfer function.

Failure to initiate any adaptive control, on the other hand, would leave unchanged the proprioceptive input, while changing both the nature of the error distribution and again, the resulting transfer function. In short, whether or not an adaptive response is implemented, the transfer function will change. If adaptation occurs, the response will change as well. If it does not, then the error distribution will be altered.

However, even provided with only two sources of information (transfer function plus error or control response) rather than three, a comparison of number of channels still favors the control mode over monitoring. Assuming that there is some degree of independence of information processing along the channels, the probability and/or speed of detecting change information along any one of two channels characterizing the control mode, should be greater than that of detecting change along the single visual channel available in the monitoring mode.

Differential sensitivity to visual vs proprioceptive information. Although a strict comparison of the number of channels of information available to a decision mechanism favors control over monitoring, an important caution should be noted. As described above, the operator is able to trade off the strength of the failure occurrence "signal" along the visual vs proprioceptive channel, to the extent that he engages in some degree of compensatory adaptation. As adaptation increases, proprioceptive "signal strength" increases at the expense of visual error "signal strength." Thus the prediction based upon the difference in number of information channels -- that control detection will be superior to monitoring detection -- is predicated upon the assumption that detection of change is equally efficient along all the channels (proprioceptive, transfer function, and visual). In other words this approach assumes that, whichever channels are employed in the control mode, their joint signal will be more easily detected than the single visual signal in the monitoring mode.

Mitigating against this conclusion, however, is a body of literature in psychology suggesting that the sensitivity to proprioceptive information is reduced relative to visual information particularly when the two sources are available at the same time and are conveying conflicting information (e.g. Jordan, 1972; Klein and Posner, 1974). Such a conflict, in fact, describes precisely the stiuation in which an operator has successfully adapted to a change in control dynamics. Under these circumstances, the visual error channel is providing information describing normal operation (since the appropriate gain, or lead-lag adjustment, has presumably been initiated to restore the original open-loop transfer characteristics), while the less sensitive kinesthetic channel conveys the information that a change has in fact been implemented. The predicted consequence of this conflict situation is that the operator will be less likely to detect the change than he would had no adaptation been achieved, the latter condition of course producing a visual signal equivalent to the monitoring mode. McDonnell (1966), in fact, has noted anecdotally such instances in which successful adaptation has been coupled with the failure to detect dynamic system changes.

Workload differences. A second characteristic of the manual control mode that predicts a reduced sensitivity to the occurrence of failures relates to the greater workload imposed by tracking than by monitoring. Numerous examples may be cited from behavioral literature that demonstrate the attention demands of purely perceptual tasks such as monitoring to be less than those of tasks such as tracking in which a requirement for the selection and execution of responses is also imposed (e.g. Keele, 1973; Kerr, 1973). This finding is verified as well in a direct comparison of controlling vs autopilot monitoring in the simulator (Johannsen, Pfender, and Stein, 1976). In the framework of the present analysis, if monitoring for and responding to failures is regarded as a "task" separate from tracking, then since the operator's attentional resources are limited, the greater workload demands imposed in the control mode than in the monitoring mode would predict poorer performance on the added "task" of failure detection in the former condition.

Whereas, workload differences make a clear prediction of detection differences in the single-task environment, this prediction is not as apparent when the performance of additional tasks is required. A common result emanating from much dual-task research is that tasks that are in themselves simpler or less loading are at the same time more vulnerable to performance decrements in a dual-task environment, as more demanding paired tasks "capture" a greater proportion of available attentional resources (Welford, 1968). In the current context, monitoring - the

simpler task - should be more vulnerable to additional dual-task requirements than controlling.

Furthermore, to the extent that failure detection while tracking is dependent upon the processing of information integral to the tracking task, then the quality of this information -- and, therefore, the quality of detection itself -- will be preserved as tracking performance is guarded in the face of competing secondary task demands. In contrast, the quality of visual information available in monitoring will be predicted by this view to deteriorate; rendering detection while monitoring more vulnerable to loading than while controlling.

Summary

The implications of the preceding theoretical analysis are complex. In summarizing, two attributes of the controlling mode may be identified that would seemingly facilitate failure detection. A greater stability of the internal model of the system, and a greater number of channels available upon which to base failure detection decisions. At the same time, the latter advantage may be mitigated to the extent that: (a) adaptation takes place reducing the strength of a visual error signal and, (b) proprioceptive sensitivity is less than visual. In comparison the monitoring mode is also characterized by two attributes that could facilitate detections: a greater "strength" of the visual signal (if adaptation by an autopilot does not take place) and a lower level of workload

Finally, it is argued that any advantage of monitoring over controlling attributable to workload differences might itself be dissipated as the competition for attentional resources is increased by imposing concurrent tasks. Clearly this interplay of factors is sufficiently complex to prohibit precise predictions concerning the superiority of one mode over the other. It does, however, facilitate a clearer identification of the nature of the failure detection task and allows predictions to be formulated concerning the differential effect of variables such as workload or control adaptation on detection performance.

In the following experiment, independent variables of participatory mode and task workload were manipulated to determine their effect on detection. Analysis techniques were then employed in an effort to identify further the nature of the processes operating in detection performance.

METHOD

Subjects

The subjects were three right-handed male university students enrolled in basic flight training courses at the Institute of Aviation. Subjects were paid at a rate of $2.50 per hour.

Apparatus

The basic experimental equipment included a 3 x 4 inch Hewlett Packard Model 1300 CRT display, a spring-centered, dual-axis tracking hand control (with an index-finger trigger) operated with the other hand, and a Raytheon 704 16-bit digital computer with 24k memory and A/D, D/A interfacing that was used both to generate inputs to the tracking display and to process responses of the subjects. The subject was seated on a chair with two arm rests, one for the tracking hand controller and one for the side-task finger controller. The subject's eyes were approximately 50 centimeters from the CRT display.

Tracking tasks. The primary pursuit-tracking task required the subject to match the position of a cursor with that of a target which followed a semi-predictable two-dimensional path across the display. The target's path was determined by the summation of two non-harmonically relate sinusoids along each axis. The frequencies were: X-axis, .08 and .05; Y-axis, .08 and .05. The position of the following cursor was controlled jointly by the subject's control response and by a band-limited forcing function with a cutoff frequency of .32 Hz for both axes. Thus the two inputs to the system were well differentiated in terms of predictability, bandwidth, and locus of effect (target vs cursor). The control dynamics of the tracking task were of the form $Y_c = \frac{1-\alpha}{s} + \frac{\alpha}{s^2}$ for each axis, where α was the variable parameter used to introduce changes in the system dynamics. These changes, or simulated failures, were introduced by step changes in the acceleration constant α from a normal value of .3, a mixed velocity and acceleration system with a high weighting on the velocity component, to $\alpha = .9$, a system that approximates pure second order dynamics.

As the loading task, the Critical Task (Jex McDonnel and Phatac, 1967), was employed. This was displayed horizontally at the bottom of the screen and required the subject to apply force to the spring-loaded finger control in a left-right direction to keep the unstable error cursor centered on the display. The value of the instability constant λ in the dynamics $Y_c = \frac{k\lambda}{s-\lambda}$ was set at a constant subcritical value. Two values ($\lambda = .05$ and $\lambda = 1.0$) were employed on different dual task trials.

Experimental Task

Subjects participated in five experimental sessions of which the first two were devoted entirely to practice on the tracking and detection tasks, and the last three used to generate the experimental data. During the first practice day the subject performed only the two-dimensional pursuit tracking task. In the manual (MA) condition the subject performed the tracking manually while in the autopilot (AU) condition, his role in the control loop was replaced by a simulated autopilot control dynamics consisting of a pure

gain and effective time-delay. The open loop gain was set at a constant value for all subjects, and the time delay value was adjusted for each subject to obtain an error measure in the AU condition equivalent to the operator's performance in the MA condition. This value of time delay was maintained throughout the rest of the experiment. Each trial, MA or AU, lasted 150 seconds.

To give the subjects some experience with the failed condition (i.e., the higher acceleration in the control dynamics), the subject received two trials (one AU and one MA) in which he tracked (or viewed the autopilot tracking) only the failed dynamics. Two demonstration trials were then presented in which the subject tracked in the regular condition, but the onset of each failure was cued by the presentation of a "F" on the screen. The subject was instructed to press the trigger to return the system to normal only upon the detection of the nature of the change. This training period was then followed by 3 regular detection trials (4 AU, 4 MA in alternating order). Each trial contained either 4 or 6 failures so that a total of 20 failures were presented in each mode.

The presentation of the failure was generated by an algorithm that assured random intervals between presentations and allowed the subject sufficient time to establish baseline tracking performance before the onset of the next change. Task logic also insured that changes would only be introduced when system error was below a criterion value. In the absence of this latter precaution, changes would sometimes introduce obvious "jumps" in cursor position.

During these detection trials, the detection decision was recorded by pressing the trigger on the control stick. This response presented a "T" on the screen and returned the system to normal operating conditions via a four-second ramp to the prefailure dynamics. If the subject failed to detect the change, the system returned to normal after six seconds. This was an interval within which it was assumed, on the basis of pretest data, that responses would correspond to detected failures and not to false alarms. The subjects were told to detect as many changes as possible as quickly as possible.

On the second day (dual-task) the subject performed the primary tracking task together with a side task, the Critical Task. After a refresher trial in the MA mode, the subject received a series of training trials to practice the side task, first in the AU and then in the MA mode. When acceptable criteria were achieved in the Critical Task and MA tracking individually, the subject then carried out these tasks together with the failure demonstrations, as described above.

Eight more experimental trials were then presented in which the subject performed all three tasks (tracking or monitoring, Critical Task, and failure detection). Two trials were presented in each mode at each level of Critical Task difficulty (λ = 0.5, 1.0). The subject was instructed to "do the side-stick task as efficiently and accurately as possible." The instructions, therefore, clearly defined the side task as the loading task while allowing performance on the tracking and detection tasks to fluctuate in response to covert changes in available attentional resources. In this manner, workload demands were experimentally manipulated, rather than being passively assessed.

During each of the final 3 days, following presentation of 4 warmup trials, the subjects received 2 replications of each of the six experimental conditions: AU and MA under single, and under the two dual task conditions. The order of presentation of the 12 experimental trials was counter-balanced across subjects and across days within a subject. The task logic, instructions, and experimental procedure was otherwise identical to that on days 1 and 2.

ANALYSIS

Assumptions from signal detection theory (Green and Swets, 1966) were employed to account for detection performance in terms not only of the proportion of failures detected (hit rate), but also the number of detection responses made in the absence of failures (false alarms). The signal detection-based sensitivity index reflects changes in both of these values. Some modification of classical signal detection analysis procedures was required because of the undefined nature of the response interval (Watson and Nichols, 1976). According to this procedure it is necessary initially to specify the interval following each failure signal to be designated as a "hit" interval. The data from a number of pretests, indicated that the distribution of subject responses, following signal occurrence, showed a peak at around three seconds and reached a relatively stable baseline by six seconds following a failure. Therefore, six-second intervals were defined as hit intervals, and the measure P(HIT) was simply the number of detection responses falling within the interval divided by the total number of intervals. The remaining duration of the trial (150 seconds - 10 x 4 or 10 x 6 depending on whether it was a 4-failure or 6-failure trial)* was similarly subdivided into six-second false alarm intervals. The measure P(FA) was computed as the number of false alarms divided by the number of false-alarm intervals.

Because of the relatively small number of signals presented, and the questionable applicability of the formal signal detection theory assumptions to the current data, the nonparametric measure of the area under the ROC curve, P(A), was employed as the bias-free measure of sensitivity (Green and Swets, 1966). Values to this measure were computed from the P(HIT) and P(FA) data by reference to tables in McNicol (1972).

This measure produces a score varying from 0 to 1.0 for which 0.5 represents chance performance and 1.0 represents perfect accuracy. Both the P(A) measure and the mean and standard deviation of detection latencies were computed at the end of each trial.

* The extra four seconds relates to the four-second ramp discussed on page

Tracking performance. Tracking measures of vector error, vector stick position, and Critical Task error were sampled every 60 msec and stored on digital tape for later data analysis. In addition, on a fourth channel, the occurrences of failures and responses were recorded. At the end of each trial, the RMS vector error and RMS error on the Critical Task (if performed) were computed.

RESULTS

Figure 2 presents the detection performance (sensitivity and accuracy data) averaged over subjects as a function of conditions, along with the data of individual subjects. The rationale for the joint speed-accuracy representation of Figure 2, is that these two variables represent different manifestations of an underlying performance metric. In any effort to compare "performance" across conditions then, the joint implications of speed and accuracy must be taken into account. For example, a condition that produces a high accuracy of responding might do so at such a prolonged latency that the utility of that decision in a real-world context is less than that of a more rapid decision with slightly lower expected accuracy. Thus the two dimensional representation can be conceptualized as containing a performance axis running from poor (slow and inaccurate) performance in the lower right to good (fast and accurate) performance in the upper left. Individual data points may be projected onto this axis. Furthermore, a speed vs. accuracy bias may be viewed as an axis orthoganal to the performance axis, and points also projected and compared along this axis. The units assigned to the performance index are clearly arbitrary but do require that an assumption be made with regard to the relative weighting of accuracy vs latency in detection. This weighting defines the scaling along the two axes or, equivalently, the slope of the performance axis. In the current data the weighting was made proportional to the intra subject variability of each measure. The elipsoids surrounding each point represent ±1 standard error confidence estimates along the latency and accuracy dimensions.

Figure 2 indicates that detection in the manual mode is apparently superior to autopilot detection. This superiority is manifest at all 3 levels of workload by a large reduction in response latency. This reduction more than compensates on the performance axis for the small loss in accuracy evident in the single task and easy dual task conditions (λ = .5). For the most difficult dual task condition, (λ = 1.0) manual detection is favored by both latency and accuracy.

A second aspect of the data of Figure 2 concerns the effect of concurrent task workload. The requirement to perform the critical task leads to a deterioration of detection performance in both modes. Increasing critical task difficulty also causes AU detection performance to decrease further, but counter intuitively appears to improve MA detection. Both trends are consistent across the data of all 3 subjects. The joint effect of dual task load and participatory mode thus takes on a form that sub-

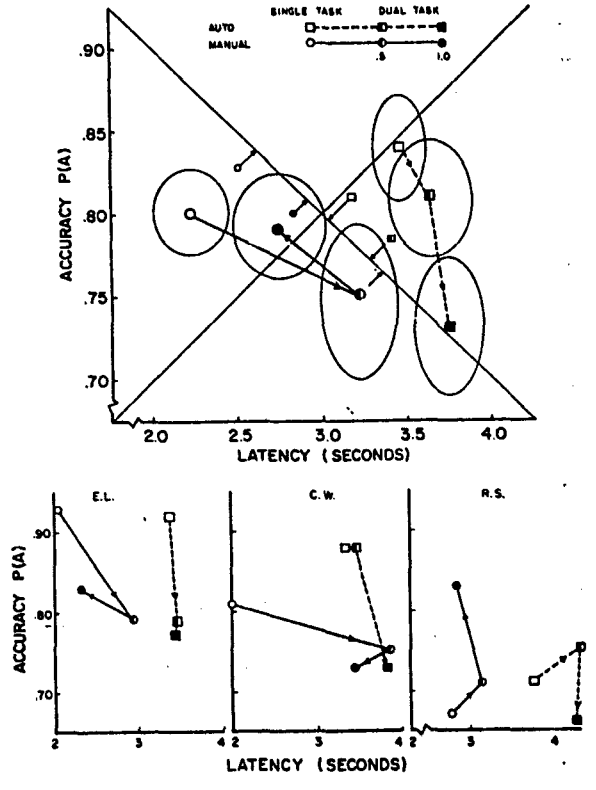


Figure 2. Detection latency and accuracy data of subject means (top) and individual subjects (bottom). One standard error confidence elipsoids surround subject data points.

stantiates data collected on 4 subjects in a shorter pilot experiment: namely that the derogatory effect of increasing concurrent task demands is greater in the AU mode than it is in the MA mode.

The data from the two tracking tasks are shown in Figure 3 presented again with the individual subject data. It is evident that under single task conditions, manual performance was closely matched by autopilot performance, and that an orderly increase in MA primary tracking error was produced by the increasing demand for processing resources imposed by the critical task requirement. This means that the improved detection in the difficult dual task MA condition cannot be attributed to improved tracking performance. Performance on the critical task itself was only slightly effected by its own difficulty, thus serving as a guarantee that this manipulation was an effective way of controlling the resources available for the primary tracking and detection tasks.

Figure 4 presents the ensemble averages of the tracking error and stick velocity following failures. Separate averages were computed for both detected and missed failures. The data are only presented for the single and the difficult dual task conditions. It may be argued that, to the extent that these profiles climb from their initial pre-failure level, there is information contained in the profiled signal that could be extracted by the subject as a cue that a failure had occured. Furthermore, to the extent that the hit and miss profiles diverge, evidences is provided that this information was different on the two classes of trials and suggests that the information was used for detection. The converse however, does not necessarily follow.

Following this line of reasoning, the data from the auto condition (4a) indicate that failure information was present in the error signal, and was utilized in detection in the single task condition and to a lesser extent in the dual task condition. The lesser degree of separation here seemingly accounts for the greatly reduced detection accuracy in the dual task condition. An interesting reversal of this trend is evident in the manual condition (4b), where the error profile is greater on missed than on detected trials. This observation suggests that in response to the failures, subjects were making an adaptive adjustment of their control strategy in order to regulate tracking error (addition of lead to compensate for the higher order dynamics). It is possible that this adjustment was initiated prior to overt detection, and the resulting proprioceptive cues were employed as a basis for the detection decision. This hypothesis is neither supported nor refuted by the ensemble averages of response velocity in Figure 4c. It is apparent that some adaptation does take place, as the profiles rise from their pre-failure level. However, the only distinction between the hit and miss profiles suggests that there may be greater, rather than less response velocity associated with miss trials.

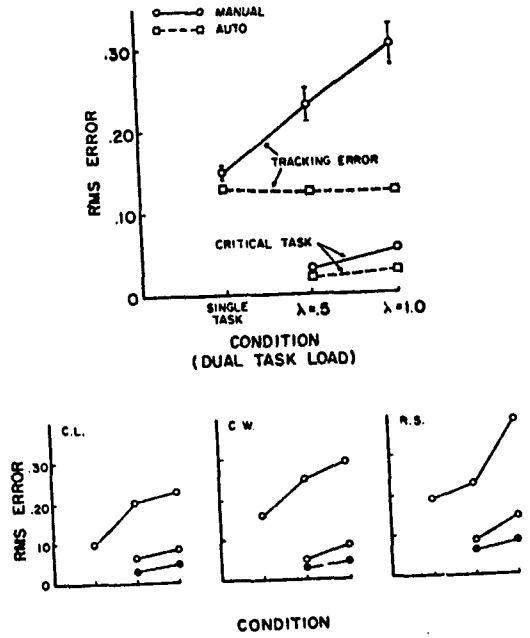


Figure 3. Primary tracking and critical task tracking erro. Subject means (top) individual subjects (bottom).

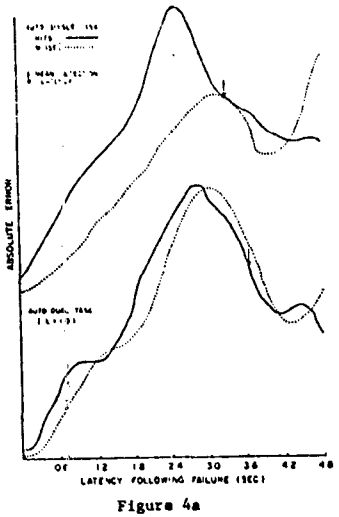


Figure 4a

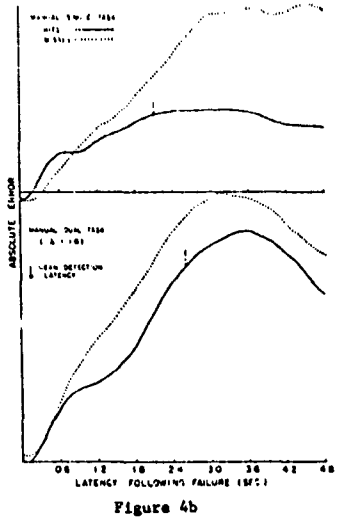


Figure 4b

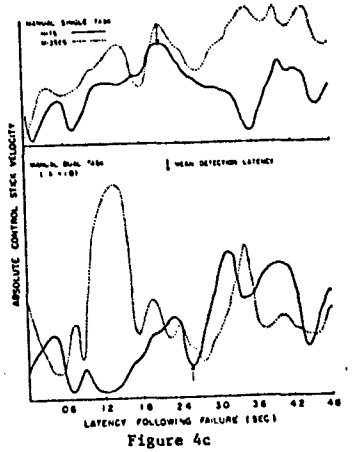


Figure 4c

Figure 4. Ensemble averages of tracking signals following failure occurrence.

In a related effort to discover the cues employed in reaching the detection decision, multiple regression analysis was used to predict detection latency, as a function of characteristics of the tracking error and response signals sampled at various time points following failure occurance. To the extent that a subject is relying upon a particular characteristic of the signal (e.g. error velocity) as an indication of failure, the value of this signal at some latency following the failure should correlate highly with detection latency. The results of this analysis failed to produce any strong trends however. In the AU condition characteristics of the error signal (error & error velocity) were low, but consistent predictors of response latency. The correlations with latency of the two error characteristics, sampled at .6, 1.2, 2.4 seconds following failure ranged from r = -.06 to -.39 (median r = -.12) for error and from r = -.10 to -.40 (median r = -.28) for error velocity. In the MA condition the hypothesis described above, that proprioceptive channels from control adaptation might serve as a detection cue, did not receive support. While this hypothesis predicts a negative correlation of response velocity (as an index of control adaptation) with latency, the obtained correlations between these variables were uniformly low and non-reliable.

DISCUSSION

The present data seemingly substantiate the conclusion drawn by Young (1969), that the sensitivity to failures is greater when the subject is an active participant in the control loop than when he is a passive monitor. While it is acknowledged that this superiority relation is undoubtably sensitive to particular task characteristics, for example the nature of the failure, the dynamics of the autopilot and its degree of adaptability, two characteristics of the current results argue for acceptance of their reliability. (1) The characteristics of the autopilot mode were clearly biased in favor of detection, in that the autopilot used in this experiment was incapable of any adaptive response which might attenuate the strength of the failure-induced error signal. This of course was not the case with the adaptive human operator in the MA mode, yet despite this advantage AU detection was found inferior. (2) One characteristic substantiates the data of all subjects that have run on earlier pretests in our laboratory.* Namely that AU detection, while not invariably less accurate, is <u>always</u> considerably slower, a point of note in considering that Young reported only latency data in describing his findings.

In speculating upon the reason for this shift along the speed-accuracy bias with participatory mode, it is possible that the greater motor involvement of the operator while engaged in active tracking responses (MA mode) serves to "prime" the response mechanism or lower the detection criterion, causing a detection response to be triggered on the basis of

*The present data represents the culmination of 4 experimental pretests. The third of these pretests, identical in format to days 1 & 2 of the current experiment generated that data that were presented at the June meeting in Cambridge.

less perceptual evidence. In the present data, while this evidence is less in the manual mode, it is presumably of equal or superior quality, thus maintaining nearly equal accuracy despite the more rapid responses.

A second noteworthy aspect of the data relates to the interaction of participatory mode and dual task load, again a consistant finding across all 3 subjects (and the previously run subjects in the pretest). This finding is in accord with the argument made in the introduction relating to the greater vulnerability to the withdrawal of attentional resources of the less demanding monitoring task. It is also a finding which, if substantiated has certain practical implications. These concern the conditions under which the greatest advantage is gained in terms of failure sensitivity by keeping the operator in the control loop -- that is, conditions of higher workload. Unfortunately these conditions are at the same time those under which other arguments might dictate replacement of the human operator by the autopilot.

With regard to the cues utilized in detection and the sources of MA superiority outlined in the introduction the data were not greatly illuminating. The greater stability of the internal model remains a viable source of superiority that is in no way contradicted by the present results. Similarly the data also suggest in the MA mode the activation of all three "failure information channels" described in Figure 1, and thus a greater number of information channels than in monitoring. Error information is clearly present (Figure 4b), while control adaptation appears to take place as well, leading to the availability of proprioceptive information. This is indicated by the reduced error profiles of the manual condition - particularly for hit trials (Figure 4b), and by the modest increase in response velocity (Figure 4c). One puzzling aspect of the data is that neither the ensemble averages nor the multiple regression analysis provided any evidence that response velocity was actually used in reaching the detection decision. It is of course possible that this particular measure was not the appropriate one to tap the process of control adaptation, and further investigation is necessary to determine alternative measures.

It is apparent that the greater workload of the MA condition does not greatly interfere with the detection process (or if it does, this interference is more than compensated by the favoring factors of model stability and information channels). The task of detection is seemingly one that is sufficiently integral with that of controlling, that a kind of dual task facilitation results when the two are performed concurrently. This is a characteristic that preserves detection sensitivity in the face of increasing dual task demands.

Finally, some mention should be made concerning the presence of individual differences. To some extent these are inevitable, particularly in a task configuration as complex as the current one, requiring dual task performance in the AU mode and triple task performance in the MA. Given the subject's flexibility to allocate resources differentially to the two or three tasks, as well as his ability to adapt various criteria on the speed-accuracy detection bias, it is perhaps somewhat surprising that the individual subject data in Figures 2 and 3 are as consistent as they are. Nevertheless, the importance is acknowledged of acquiring more data to replicate and substantiate the trends reported here.

Curry, R. E., and Gai, E. G. Detection of random process failures by human monitors. In T. B. Sheridan and G. Johannsen (Eds.), Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Ephrath, A. R. Detection of system failures in multi-axes tasks. Proceedings of the 11th annual NASA-University conference on manual control. NASA TMX, 1975, 62, 464.

Gould, E. E., and Fu, K. S. Adaptive model of one human operator in a time-varying control task. Proceedings of the second annual NASA-University conference on manual control, NASA SP, 1966, 128, 65-96

Green, D. M., and Swets, J. A. Signal detection theory and psychophysics, New York: John Wiley and Sons, Inc., 1966.

Hays, W. L. Statistics. London: Holt, Rinehart, and Winston, 1969.

Jex, R., McDonnell, J. D., and Phatak, A. V. A "critical" tracking task for manual control research. IEEE Transactions on Human Factors in Electronics, 1966, 7(4).

Johannsen, G., Pfendler, C., and Stein, W. Human performance and workload in simulated landing-appraoches with autopilot-failures. In T. B. Sheridan and G. Johannsen, (Eds.), Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Jordan, T. C. Characteristics of visual and proprioceptive response times in the learning of a motor skill. Quarterly Journal of Experimental Psychology, 1972, 24, 536-543

Keele, S. W. Attention and human performance. California: Goodyear Publishing Company, 1973.

Kerr, B. Processing demands during mental operations. Memory and Cognition, 1973, 1, 401-412.

Klein, R. M., and Posner, M. I. Attention to visual and kinesthetic components of skills. Brain Research, 1974, 71, 401-411.

McNicol, D. Signal Detection Theory. London: George Allen & Unwin Ltd. 1972

McDonnell, J. D. A preliminary study of human operator behavior following step changes in the controlled element. IEEE Transactions on Human Factors in Electronics, HFE Tech. 1966, 125-129.

Miller, D. C., and Elkind, J. J. The adaptive response of the human controller to sudden changes in controlled process dynamics. IEEE Transactions Human Factors in Electronics, HFE, 1967, 8 (3), 218-223.

Pew, R. W. Human perceptual-motor performance. In B. H. Kantowitz, (Ed.), Human information processing tutoriale in performance and cognition. New York: Wiley, 1974.

Sheridan, T. B. Preview of models of the human monitor/supervisor. In T. B. Sheridan and G. Johannsen, (Eds.), Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Sheridan, T. B., and Johannsen, G. Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Swets, J. A., and Kristofferson, A. B. Attention. Annual Review of Psychology, 1970, 21, 339-366.

Taylor, M. M. Detectability theory and the interpretation of vigilance data. Acta Psychologica, 1967, 27, 390-399.

Veldenhysen, W., and Stassen, H. G. The internal model: What does it mean in human control. In T. B. Sheridan and G. Johannsen, (Eds.), Monitoring Behavior and Supervisory Control, New York: Plenum Press, 1976.

Watson, C. S., and Nichols, T. L. Detectability of auditory signals presented without defined observation intervals. Journal of Acoustic Society of America, 1976, 59, (3).

Welford, A. T. Fundamentals of Skill. London: Metheun, 1968.

Young, L. R. On adaptive manual control. IEEE Transactions on Man-Machine Systems, V. MMS-10, 1969, 4, 292-331.

PREDICTION OF PILOT RESERVE ATTENTION CAPACITY
DURING AIR-TO-AIR TARGET TRACKING

E. D. Onstott
W. H. Faulkner

Northrop Corporation
Aircraft Group
Hawthorne, California

Proceedings of the
Thirteenth Annual Conference
On Manual Control
June 15-17, 1977

Massachusetts Institute of Technology
Cambridge, Massachusetts

PREDICTION OF PILOT RESERVE ATTENTION CAPACITY
DURING AIR-TO-AIR TARGET TRACKING

E. D. Onstott
W. H. Faulkner

Northrop Corporation
Aircraft Group
Hawthorne, California

ABSTRACT

Reserve attention capacity of a pilot has been calculated using the Northrop pilot model that allocates exclusive model attention according to the ranking of task urgency functions whose variables are tracking error and error rate. The modeled task consisted of tracking a maneuvering target aircraft both vertically and horizontally, and when possible, performing a diverting side task which was simulated by the precise positioning of an electrical stylus and modeled as a task of constant urgency in the attention allocation algorithm. The urgency of the single loop vertical task is simply the magnitude of the vertical tracking error, while the multiloop horizontal task requires a nonlinear urgency measure of error and error rate terms. Comparison of model results with flight simulation data verified the computed model statistics of tracking error of both axes, lateral and longitudinal stick amplitude and rate, and side task episodes. Full data for the simulation tracking statistics as well as the explicit equations and structure of the urgency function multi-axis pilot model are presented.

INTRODUCTION

Historically, the development of mathematical models of the human controller has been based upon linear continuous operators, such as the Laplace transform. This formulation lends itself to ready analysis of piloted system stability and frequency response for a very restricted class of problems, but the progress made using this approach had led to the development of pilot-vehicle modeling as a widely employed and established discipline. A large bibliography now exists, and the dynamic properties of single-axis time-invariant human control are recorded in detail. If one considers the problem of continuous linear systems driven by Gaussian processes, additional performance information can be obtained since the input and output power spectra are related. For a zero mean process, the integral of the output spectrum is simply the mean square of the time history. In this way, time domain information can be obtained for problems where root mean square performance is useful to optimize.

Although linear analysis of time-invariant problems has been useful for design and evaluation applications, it has not led to sufficiently comprehensive studies of large scale problems such as weapon delivery and loss of control at high angles of attack. In order to study such examples, the total system model must contain the following features:

- Nonlinear and time-varying aircraft dynamics.
- Nonlinear and time-varying pilot dynamics.
- Multi-axis as well as multi-loop pilot operation including exact modeling of time sharing, sampling, changing tracking criteria, and threshold effects.
- Predictive model adjustment using optimization.
- Easily applied computational methods which permit the modeling of large systems including gross nonlinearities.

In order to meet these requirements, Northrop has developed an approach that is based on the following two techniques:

- Digital simulation as a piloted handling qualities analysis method.
- Time domain pilot modeling which includes decision algorithms for multi-axis control and instrument scanning, and a direct method for modeling the pilot as shown in Figure 1.

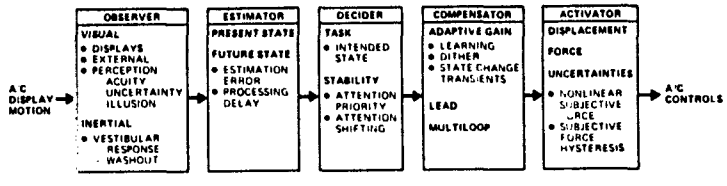


Figure 1. Total Pilot Model

These five areas of observation, estimation, decision, compensation, and actuation all contribute significantly to the total pilot-aircraft interface. Since only the compensation block lends itself to linear analysis, describing function and optimal control approaches to pilot modeling have been mainly confined to this area. The other functions have been traditionally relegated to a lumped method of degrading the model performance to emulate deteriorated tracking scores through the injection of noise to the model's output. The Northrop time domain pilot model, on the other hand, allows all relevant nonlinearities and time dependent functions of the pilot to be directly represented. Estimation and observation includes both the evaluation of perception position and rate uncertainties and thresholds. Adaptive algorithms to represent the gain optimizing capability of the pilot producing the control compensation are also being developed.

There are a number of advantages in modeling pilots in the time-domain. Since frame lapsing can be used to give exact representations of transport delay, Padé approximation can be avoided. This alone improves the accuracy of statistical predictions and eliminates otherwise unpleasant transient effects. In addition to the advantages of improved accuracy, time domain models can incorporate information that is fixed in time, such as criterion governing the abandoning of one task for another. This kind of pilot behavior is crucially important in multi-task missions and since the decision points are functions of the system state variables it cannot be modeled by either frequency domain or state space methods.

There have been three main approaches previously taken in attempts to extend single axis model theory to multi-axis tasks. All of these recognize that the human must operate as a time shared device when faced with difficult control tasks. This intermittent operation degrades the performance of each task from what the pilot would achieve in continuous control. As might be expected, these three approaches are (1) decrease the model gain from the optimum for continuous control, (2) increase the time delay to account for the periods of inattention, and (3) inject filtered noise to imitate the spectral content of the shifting pilot control.

The problem with these three approaches is this: The human pilot is quite discriminating about when he will abandon the control of one task to take over the control of another. This leads to a sampling criterion that is functionally dependent on the total system variables. In no way can this be regarded as a purely random, or a regular sampling. Thus a multi-axis pilot model must contain an algorithm that determines when control shifting takes place, and the model must be computed in a way that preserves this information.

By using the method of digital simulation, the exact functional criterion, by which a pilot decides his control, can be directly computed without the gross distortions of linearization. The development of the form of these urgency criteria has now advanced

to the point where they can be determined from 1) the system dynamics, 2) the task, and 3) the appropriate human factor information about the pilot.

Let x_1 be the state variables of one axis, x, of a two axis task, and let the other axis, y, be represented by y_1. Then the attention allocation criterion for the x axis is satisfied identically with the inequality

$$U_x(x_1) \geq U_y(y_1) \tag{1}$$

where U_x and U_y are the urgency functions of the x and the y tasks. These functions are always nonlinear in the state variables, but fall into several precise classes. Some of these classes have been well explored and Reference 1 considers the dependency of the urgency functions on error rate information.

The multi-axis urgency function model thus consists of simple linear pilot dynamics, along with the control criterion of (1). Whichever axis has the larger urgency function gets the corrective control attention. The adjustment of the linear coefficients can usually be obtained by an easy search starting with the optimum single axis coefficients.

The success of this approach has been reported in References 2 and 3 for VTOL hover and two-axis attitude stabilization tasks in turbulence. Recently the model has been applied to the problem of air-to-air target tracking along with a third attention diversion task to measure reserve pilot attention.

TARGET TRACKING WITH VISUAL DELAYS AND SIDE TASK

In 1975, M. J. Queijo and D. R. Riley of NASA LRC performed a flight simulation study to determine the effect of time delays in visual cues on pilot tracking performance, as reported in Reference 4. The subjects controlled a five degree of freedom aircraft tracking a target that maneuvered by slow altitude oscillations in the vertical plane. By delaying the visual CRT display, evaluation of flight simulator time delays was obtained. In addition to vertical and horizontal tracking statistics, workload information was obtained by use of a side task which consisted of using an electrical stylus to tap alternately on two electrodes separated by a barrier and strapped to the subject's leg. The general availability and completeness of the reported experiment make this problem a useful one for the demonstration of the Northrop urgency decision pilot model.

The specific piloted tasks were as follows:

1) Track the target vertically. The target oscillated at a frequency of .21 radian/sec with an amplitude of ±100 feet at a distance of 600 feet ahead of the tracking aircraft.

2) Track the target horizontally. The target did not oscillate horizontally; inadvertent pilot input provided the lateral task.

3) Whenever possible, perform the side task of tapping the electrodes strapped to the leg. This tapping rate was postulated to measure pilot reserve attention capacity for the target tracking task.

In order to establish the dynamic form of the multi-axis pilot model, the following assumptions are made:

1) The pilot tracks vertically and horizontally, not in azimuth and elevation.

2) The tasks vertical tracking, horizontal tracking, and side task are performed one at a time depending on the relative urgencies U_V, U_H, U_{ST} of these tasks.

3) The side task represents a constant urgency diversion from the vertical and horizontal tasks:

$$U_{ST} = \text{Constant}$$

These pilot model assumptions can then be implemented by programming a time domain digital simulation of the pilot compensations for each task along with the urgency functions and their associated decision logic. Figure 2 shows a diagram of the complete simulation model.

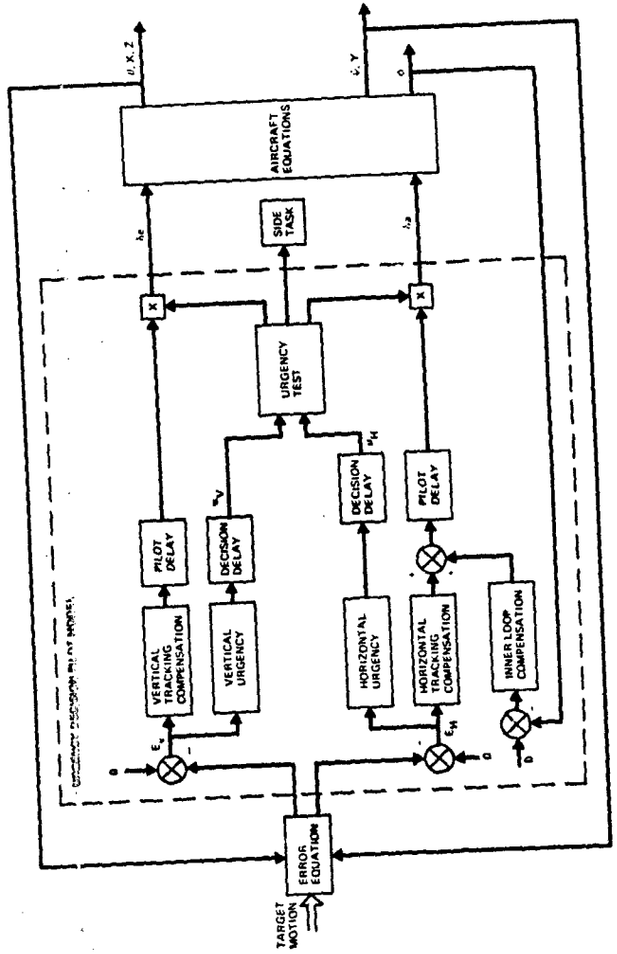


Figure 2. Control Configuration for Target Tracking Task

Consider the vertical task first. Here the tracking control is essentially to point the aircraft, so that this pilot closure is a single loop pitch tracking task. Now if ϵ, K, and T_L denote tracking error, pilot model gain, and pilot model lead, with the subscripts V and H denoting the vertical and horizontal tasks, the fixed form compensation of the vertical task can be written:

$$\begin{cases} \delta e = \text{Delay } (\tau) \left| K_V \left(\epsilon_V + T_{L_V} \dot{\epsilon}_V \right) \right| \\ \delta a = 0 \end{cases} \tag{2}$$

The associated urgency function of this single loop control task is then dictated by the general formulation of the urgency decision pilot model to be of the form:

$$U_V = \left| \alpha_V |\epsilon_V| + \beta_V \frac{\epsilon_V \dot{\epsilon}_V}{|\epsilon_V|} \right|$$

When the model is in horizontal control, it is required to track the target through changes in the heading of the tracking aircraft. This multiloop task is modeled through an inner loop attitude stabilization and an outer loop heading command tracking closure. These take the forms:

$$\begin{aligned} \delta a_1 &= \text{Delay } (\tau) \left| K_\phi \left(\phi + T_{L_\phi} \dot{\phi} \right) \right| \\ \delta a_2 &= \text{Delay } (\tau) \left| K_H \left(\epsilon_H + T_{L_H} \dot{\epsilon}_H \right) \right| \end{aligned} \tag{3}$$

$$\begin{cases} \delta a = \delta a_1 + \delta a_2 \\ \delta e = 0 \end{cases}$$

The horizontal urgency function takes the form:

$$U_H = \left| \alpha_H |\epsilon_H| + \beta_H \frac{\epsilon_H \dot{\epsilon}_H}{|\epsilon_H|} \right| \tag{4}$$

Of the various pilot model parameters, only four can be assigned typical values prior to optimization of the model performance. These values are given in Figure 3.

$\tau = 0.3$ second

$T_{L_V} = 0.5$ second

$T_{L_\phi} = 0.5$ second

$T_{L_H} = 1.5$ second

Figure 3. Pre-assigned Pilot Model Parameters for Target Tracking Task.

In order to demonstrate the predictive capabilities of the model, it is necessary to make clear the procedure for obtaining the other model parameters by optimization, and the method for calibrating the side task urgency. There were three steps involved in doing this:

1) Optimize by a gradient method the quantities

$$K_V,\ \alpha_V,\ \beta_V,\ K_\phi,\ K_H,\ \alpha_H,\ \beta_H$$

for minimum target miss distance

$$\epsilon = \sqrt{\epsilon_V^2 + \epsilon_H^2} \tag{5}$$

using no side task,

$$U_{ST} = 0 \tag{6}$$

2) Using these optimized values, vary U_{ST} until ϵ_V matches ϵ_V reported for the minimum visual delay reported for one simulation test datum.

3) Holding all quantities including U_{ST} constant for each simulation test case, vary the visual time delay by retarding the pilot model input to obtain model data.

It should be clear from the above description that the model was adjusted by optimization, with only one statistic of one test case matched to the simulation data for the calibration of the side task urgency, a quantity that could not be extracted from the Queijo - Riley report.

Two test cases, 5 and 6, were investigated by using the method as described above. Each unit of visual delay equaled 0.03125 second. Figure 4 shows the comparison data for pilot model and flight simulation for a visual delay of 1.5 units. It can be seen that not only do the tracking errors agree, but the average stick amplitude and stick rate statistics as well. Figure 5 presents the data for a visual delay of 6.5 units.

	CASE 5		CASE 6	
	Simulation	Model	Simulation	Model
ℓ_V	3.53*	3.42	3.72	3.60
ℓ_H	1.94	1.97	2.68	2.06
$\overline{\delta a}$	0.029	0.030	0.049	0.042
$\overline{\delta e}$	0.0057	0.0044	0.0089	0.0097
$\overline{\dot{\delta a}}$	0.093	0.26	0.154	0.36
$\overline{\dot{\delta e}}$	0.025	0.017	0.035	0.047

*matched data point for side task urgency calibration

Figure 4. Numerical Data for Visual Delay of 1.5 Units

	CASE 5		CASE 6	
	Simulation	Model	Simulation	Model
ℓ_V	4.09	4.29	5.05	5.82
ℓ_H	2.39	2.09	3.45	2.32
$\overline{\delta a}$	0.035	0.032	0.061	0.051
$\overline{\delta e}$	0.0067	0.006	0.011	0.02
$\overline{\dot{\delta a}}$	0.098	0.26	0.169	0.375
$\overline{\dot{\delta e}}$	0.027	0.021	0.038	0.09

Figure 5. Numerical Data for Visual Delay of 6.5 Units

Additional data are presented in graphical form for case 5 in Figures 6 and 7 which present the pilot model data for the vertical and horizontal tracking errors in comparison with ± 1 standard deviation simulation data. Pilot model data are presented in Table 1.

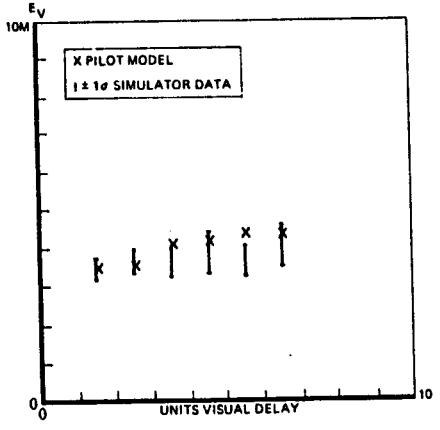


Figure 6. Case 5 Vertical Tracking Data

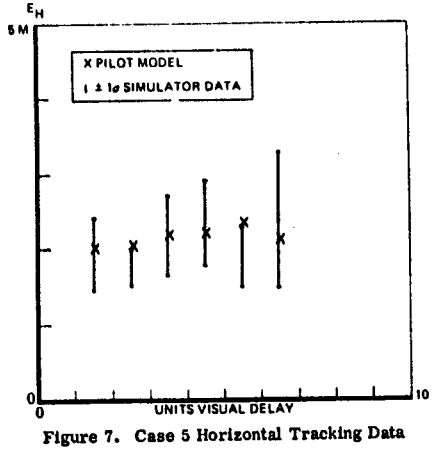


Figure 7. Case 5 Horizontal Tracking Data

Table 1. Target Tracking Pilot Model Data

	Case 5	Case 6
K_v	-0.002	-0.004
α_v	2.0	2.0
β_v	0.1	0.1
K_H	0.015	0.02
α_H	3.0	3.0
β_H	10.0	10.0
K_ϕ	0.6	0.6
U_{st}	5.5	5.5
Initial ϵ_H	0.5 m	0.5 m

A combined comparison of cases 5 and 6 vertical and horizontal tracking errors is presented in Figure 8 for the data furnished in Figures 4 and 5.

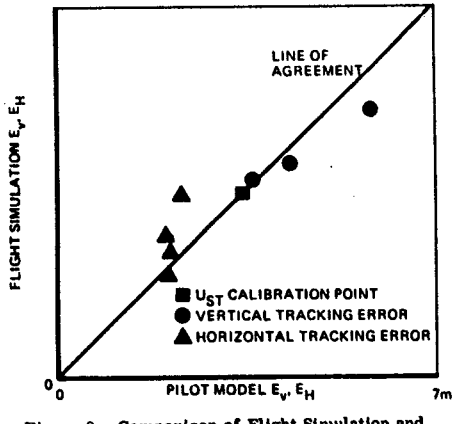


Figure 8. Comparison of Flight Simulation and Pilot Model Target Tracking Errors.

It is useful to examine a plot of tracking error versus time as it would be viewed in the sight by the pilot. Figure 9 was obtained from the pilot model, with the pilot model control episodes shown by symbols as indicated.

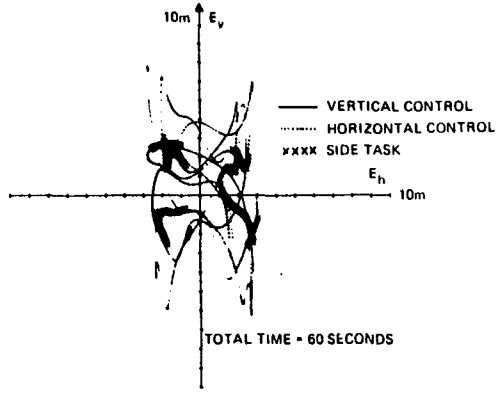


Figure 9. Time History of Pilot Model Control Episodes.

There are two important observations that can be made concerning the side task. By examination of many time histories such as Figure 9, it was clear that the side task was performed by the model only when the tracking error was less than four meters. Since the simulation pilots were reported not to perform side tasks unless the error was less than one wing span of the target, this model side task behavior is consistent with the flight simulation. The second point of comparison concerns the frequency of side task episodes. Strip chart data from the flight simulation reported in Reference 4 show that side task counts tend to occur in pairs. If the side task counting rate is halved on the assumption that the pilot usually gets the second count once he has looked down to perform the first, the counting rate is approximately equal to the frequency of side task episodes produced by the urgency decision pilot model. It should be noted that the side task has significant influence on vertical and horizontal tracking errors; since these statistics compared well, the assumption of a constant urgency model for the electrode tapping side task appears to be justified.

CONCLUSIONS

Although the data presented here has been limited to cases 5 and 6 of Reference 4, the following four conclusions can be drawn concerning the ability of the urgency decision pilot model to represent piloted target tracking:

1) The model reproduced the vertical tracking errors accurately.

2) Even though there was no horizontal target motion, the attention diversion in the pilot and in the pilot model led to similar horizontal tracking errors.

3) The assumption that the electrode tapping side task used in the flight simulation could be represented as a constant urgency task was justified by a) the vertical and horizontal tracking errors, b) the occurance of side task episodes only for tracking errors less than 4 meters, and c) the frequency of side task episodes.

These results justify the last conclusion:

4) The urgency decision pilot model can be used to predict tracking error performance and pilot reserve attention capacity for maneuvering targets by straightforward application of the fixed form model adjusted through optimization and side task urgency calibration involving only one data point.

REFERENCES

1. Faulkner, W. H., and Onstott, E. D., Error Rate Information in Attention Allocation Pilot Models, Proceedings of the Thirteenth Annual Conference on Manual Control, MIT, June 1977.

2. Onstott, E. D., Task Interference in Multi-Axis Aircraft Stabilization, Proceedings of the Twelfth Annual Conference on Manual Control, University of Illinois, May 1976.

3. Onstott, E. D., Multi-Axis Pilot-Vehicle Dynamics, Proceedings of the Tenth Annual Conference on Manual Control, AIFT/AFFDL, April 1974.

4. Quelio, M. J., and Riley, D. R., Fixed-Base Simulator Study of the Effect of Time Delays in Visual Cues on Pilot Tracking Performance, NASA TN D-8001, October 1976.

REDUCED MENTAL CAPACITY AND BEHAVIOR OF A RIDER OF A BICYCLE SIMULATOR UNDER ALCOHOL STRESS OR UNDER DUAL TASK LOAD

Mathijs Soede
Netherlands Institute for Preventive Medicine
Health Organization TNO
Wassenaarseweg 56
Leiden Netherlands

In mental load studies most research is done with tasks requiring much attention. Concurrent tasks or drugs will decrement the performance in those situations immediately. It is expected that different results will be obtained in tasks performed as an overlearned reflex.
In previously described bicycle simulator experiments both aspects are combined, i.e., the course-following task is supposed to be an attention demanding task and the balancing of the bicycle is supposed to be an over learned routine task for most Dutch people.
Experiments were carried out on this simulator with alcohol administration and a binary choice task in separate sessions, intending to reduce the subject's mental capacity. Before and after such sessions a visual evoked response measurement was done. The subject's performance was analyzed with describing function techniques. Details of this technique using two forcing functions to analyze the two dimensional task are given in other papers (Van Lunteren). The results indicate that the alcohol affects the course-following task as well as the balancing task: i.e., a general effect. The binary choice task is more specifically influencing the course-following task. The dual task shows a more pronounced effect on the recovery of the evoked response. The alcohol is delaying the recovery curve of the evoked response. A tentative explanation can be given which agrees with the performance data.

REDUCED MENTAL CAPACITY AND THE BEHAVIOR OF A RIDER OF A BICYCLE SIMULATOR UNDER ALCOHOL STRESS OR UNDER DUAL TASK LOAD

Mathijs Soede

Netherlands Institute for Preventive Medicine
Wassenaarseweg 56
Leiden
Netherlands

0 SUMMARY

In mental load studies most research is done with tasks requiring much attention. Concurrent tasks or drugs will decrement the performance in those situations immediately. It is expected that different results will be obtained in tasks performed as an overlearned reflex.
In previously described bicycle simulator experiments both aspects are combined, i.e., the course following task is supposed to be an attention demanding task and the balancing of the bicycle is supposed to be an overlearned routine task for most Dutch people.
Experiments were carried out on this simulator with either alcohol administration or a binary choice task in separate sessions, intending to reduce the subject's mental capacity. Before and after such sessions a visual evoked response measurement was done. The subject's performance was analyzed with describing function techniques. Details of this technique using two forcing functions to analyze the two dimensional simulator task are given in other papers (Van Lunteren). The results indicate that the alcohol affects the course following task as well as the balancing task: i.e. a general effect. The binary choice task is more specifically influencing the course following task. The dual task shows a more pronounced effect on the recovery of the evoked response.
The alcohol is delaying the recovery curve of the evoked response. A tentative explanation can be given which agrees with the performance data.

1 INTRODUCTION

Most of the research in mental load and mental capacity is done with the aim of protecting the subject from overloading whichcan result in a decrement of performance or an increase in errors. An important concept in this research is the reserve capacity. Some methods are known to estimate this reserve capacity; although none of these methods is enough reliable and valid, under certain conditions it is possible to give rough estimates of the mental load or its complement,the reserve capacity (Ref.1).
Research done with concurrent tasks, is in general set up from the viewpoint of strategies for allocation of attention. In this case there is no overloading of the subject and the investigator is not interested in the level of load imposed upon the subject.
Another application of a concurrent task is the use of a distraction task. In this case the presumption of a limited mental capacity is based on the fact that the primary task performance can be altered or disturbed by the

distraction task.
It is not always possible to explain the changed performance or the errors in task performance by a single channel theory (Ref.2).
It is more acceptable that the handling of information, i.e. performance of tasks by the human operator can be done in two different ways, either in an automatic way, based on routine, or in a concious way requiring much attention. This means, respectively, a performance produced by hierchical lower control systems or at the level of the central nervous system.
In this paper a pilot experiment is described in which a task is considered to be performed using control at various levels of the nervous system. This task is the stabilization of, and course following on a bicycle simulator respectively considered as a routine aspect and an attention demanding aspect of riding a bicycle. The presumption that reduction of the subject's mental capacity (i.e. at the level of the central nervous system) will cause altered performance is the basic issue of this paper.
Different causes of reduced capacity occur in practice, for instance illness, distraction, age, drugs and physical handicaps requiring extra attention to compensate for the handicap.
Two different sources of reduction are given in the experiments described in this paper i.e. the performance of a binary choice task as a distraction-task and the administration of alcohol supposing to give, among other effects, a reduction of the mental capacity.
The purpose of this research can be formulated as follows:

- To what extent is the reduction in capacity compensated for by altered strategies in task performance.
- What are the similarities and differences between capacity reduction by distraction (binary choices) and by a specific drug i.e. alcohol.

Besides these purposes the application of describing function techniques in a two dimensional cross coupled control task was of interest.
A frequently proposed method of estimating mental load is based on physiological measures like heart rate variability, blood pressure, galvanic skin response and EEG. Therefore in the reported experiments some of these variables have been measured.
Before reporting the experiments, section 2 will give some general remarks about capacity and capacity reduction to give more insight in problems connected with this type of research.

2 ASPECTS OF CAPACITY AND CAPACITY REDUCTION

In literature the concept of mental capacity is defined in different ways. Firstly there is no agreement about the unit for mental activity. Examples are the bit of information or the element of a task.
Secondly it appears that different definitions of capacity are given due to insufficient knowledge of dynamic and stochastic aspects of mental capacity.
A frequently used, operational definition of capacity is one whereby the task load is expressed as the number of choices per unit of time.
In this case the maximum capacity is sometimes defined as the maximum task load (number of choices), which can be maintained for a certain duration (Ref.3). A graph of the task load at different durations is given in Fig.1.
The momentary capacity is not equal to the capacity as given by the definition mentioned above and illustrated in Fig. 1.

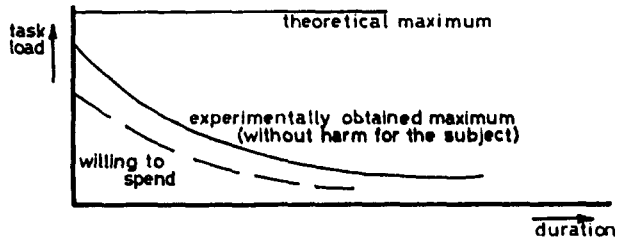


Fig. 1 The relation between a constant task load and duration of performance.

The dynamic aspect of mental capacity implies that the momentary capacity is dependent on the previously used capacity and recovery characteristics of the subject. There is not enough knowledge about this aspect to propose a model for this phenomenon.
The stochastic aspects of capacity are the fluctuations in capacity due to, for instance, variations in external influences, momentary condition and motivation of the human being.
This means that differences in capacity have to be expected between different persons and at different moments for the same person.
Based on these considerations the following has to be stated with respect to those experiments where capacity is an important notion.

- One has to control the duration of the test runs accurately to prevent different effects from the dynamic aspects of capacity.
- For the same reason the rest periods between test runs has to be controlled.
- Randomization of the sequence of different experimental conditions is advised to account for the effects of the stochastic aspects of capacity.
- More complicated is the situation when the subject's capacity is not filled up with the task load, this means a difference between the capacity held available by the subject and the capacity used for performing the task.

The decision to acquire some physiological measures in the reported experiments is based on the last remark because these measures are supposed to be related with the amount of available capacity (Ref 4,5,6).

3 DESCRIPTION OF EXPERIMENTS

3.1 THE BICYCLE SIMULATOR

An extensive description of the bicycle simulator is given in earlier publications (Ref.7). The main parts of the simulator are the following (Fig.2).

- Bicycle frame with roll possibility and rotating handle bar.
- Torsion motor to stabilize the bicycle frame
- Display system to indicate the simulated course; this system has been built with a rotating projector behind the bicycle frame and a projection screen at a distance of 2 meters in front of the bicycle frame.

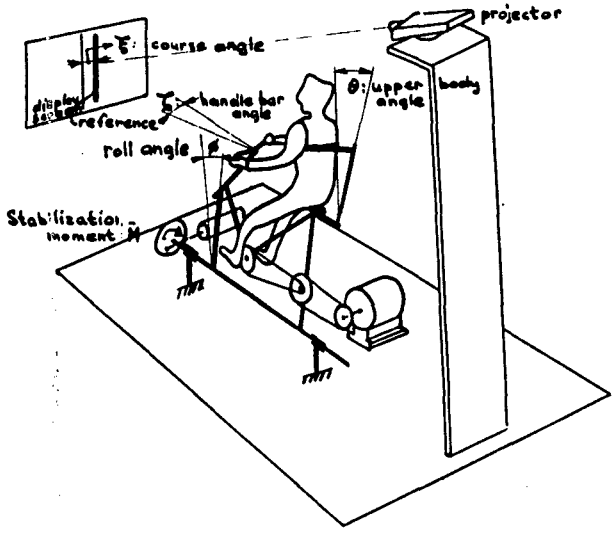


Fig. 2 A schematic drawing of the bicycle simulator and some important signals

The course is given by a vertical projected line on the screen which has a fixed reference line in the middle.

- Electric and electronic equipment to simulate forward speed, to control the torsion motor and to control the turn table of the projector.

The inputs for this system are frame roll angle, handle bar angle and a preset value indicating the forward speed (the simulator is operated as a moped).

Provisions are made to measure upper body angle and other important signals.

A block diagram is given in Fig.3. This block diagram also presents the elements of the describing function of the human operator: i.e. the transferfunctions $H_1 \ddagger H_4$.

It is clear from this block diagram that the cross coupling of elements in the simulator give the possibility to stabilize with handle bar rotation or with upper body movements. This gives rise to the element H_2 in the

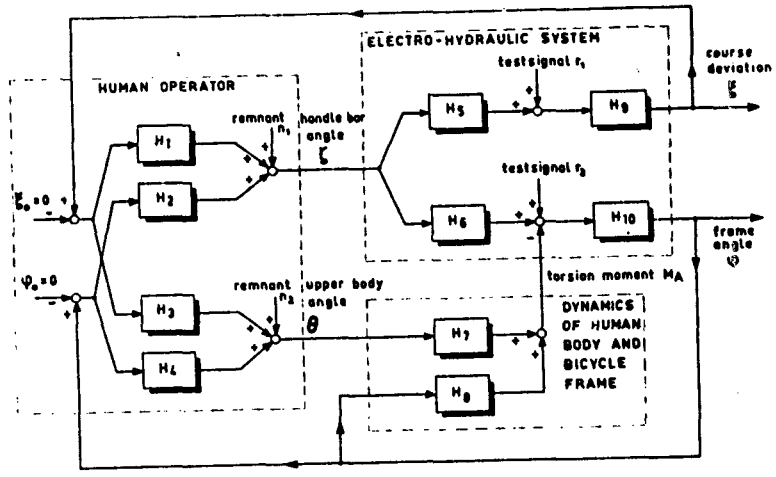


Fig. 3 A block diagram of the man-bicycle system including a cross coupled course following system and a stabilization system.

describing function. It is not possible to follow the course by upper body motions. However, interference in body movements and arm movements resulting in handle bar rotations will probably cause a relation between roll angle and handle bar angle. Therefore the element H_3 is proposed to describe this possible relation. The forcing functions r_1 and r_2 simulate respectively a course and the effect of side-wind on the man-bicycle system. The simulator gives a reasonable resemblance with the real situation. The missing of wind and longitudinal movement are not very important. The different direction of the accelleration on the human body gave some troubles in the first training sessions, however the trained subjects were not troubled with this phenomenon.

3.2 SET-UP OF THE EXPERIMENTS

The experiments consisted of three series of measurements each during a morning from 8.00 a.m. to 1.00 p.m.. At the first morning an additional binary choice task was given to the subject performing the bicycle task. The second morning was appointed for bicycle riding with alcohol. The last morning was intended to be a reference measurement to compensate for

diurnal rhythms.
The division of the day is given in Fig. 4 and shows that four task measurements were made each day.

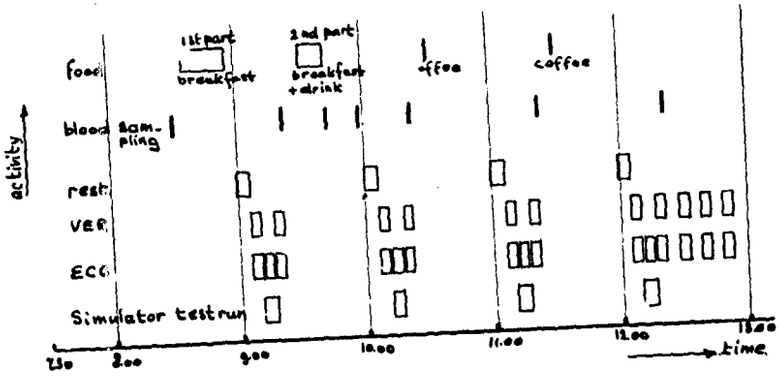


Fig. 4 The time schedule of a day of experiments.

The first measurement on the day was always without distraction task or alcohol.
The electrocardiogram was measured before, during and immediately after a simulator test run.
The visual evoked response was measured together with the corresponding ECG measurements only before and after task performance. The latter measurements were taken in a seperate EEG-room.
The subjects were 6 male students. A medical examination was carried out beforehand. The subjects were asked not to use alcoholic drinks during the week of the experiments. The training on the bicycle simulator and on the binary choice task consisted of six hours divided over a three days period. All subjects had much experience with normal bicycle riding.

3.3 THE BINARY CHOICE TASK

Distraction was caused by a binary choice task (Ref.8). The task stimulus consisted of a randomly generated high or low tone, presented with headset to the subjects. These tones had to be answered by operating a microswitch with respectively the right and left hand. The microswitches were mounted in a convenient place on the handle bar of the bicycle.
The number of choices given at the first day of experiments was fixed at 0,45,30 and 15 per minute in the first, second, third and last test run on that day respectively.

3.4 THE ALCOHOL ADMINISTRATION

At the second day of experiments the subject was given a drink with alcohol resulting in blood alcohol concentrations as given in Fig. 5.

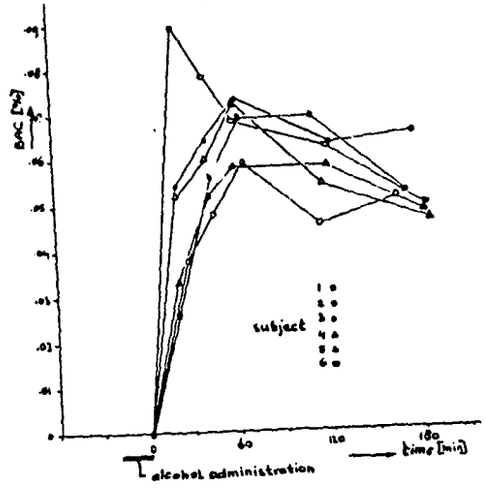


Fig. 5 The individual blood alcohol curves.

The alcohol administration occurred just after the first test run.
The blood alcohol concentrations were determined by means of venous punctures at selected moments between the test runs. To account for emotional effects caused by the venous puncture, blood samples were taken each day.

4 METHODS OF ANALYSIS

4.1 ANALYSIS OF PERFORMANCE ON THE BICYCLE SIMULATOR

The model chosen for the behavior of the rider of the simulator consists of four elements i.e. the describing functions of the relations between the following signals:
- H_1: course angle ζ and handle bar angle ζ
- H_2: roll angle ϕ and handle bar, angle ζ
- H_3: course angle ζ and upperbody angle θ
- H_4: roll angle ϕ and upperbody angle θ

The technique for estimating the parameters in this model is given elsewhere (Ref. 9, 10), therefore only a summary is given here.
The two forcing functions used, consist each of a sum of ten sinusoïds, having a flat amplitude spectrum and an overall repetition time of 4 minutes. Due to the necessity of decoupling the signals in the system with respect to each of the forcing functions, the frequencies in each of the forcing functions have to be different. Therefore the frequencies are chosen in pairs closely together. The situation shown in Fig. 6 is the situation with decoupled signals i.e. $\phi^*(\nu)$, $\xi^*(\nu)$, $\zeta^*(\nu)$, and $\theta^*(\nu)$. (ν is the frequency in Hz).

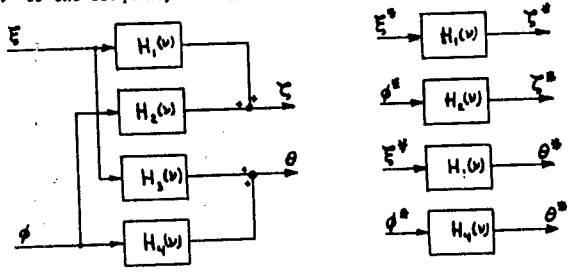


Fig. 6 The describing function model of the human operator before and after decoupling of the signals.

Some derivation leads to the following expression:

$$\zeta^*(\nu)=\zeta_1(\nu)-\frac{\zeta_2(\nu)}{\phi_2(\nu)}\,\phi_1(\nu) \qquad (1)$$

The signals are split up with respect to the forcing functions $r_1(\nu)$ and $r_2(\nu)$. The indices in Eq.(1) relate to the indices of the forcing functions. At a frequency ν_k of a sinus component of test signal $r_1(\nu)$, it turns out that $\zeta_2(\nu_k)=\phi_2(\nu_k)=0$ at that particular frequency.
Assuming that the transfer functions are sufficiently smooth, it can be stated that if $\Delta\nu$ is small:

$$\frac{\zeta_2(\nu_k)}{\phi_2(\nu_k)}\approx\frac{\zeta_2(\nu_k+\Delta\nu_k)}{\phi_2(\nu_k+\Delta\nu_k)} \qquad (2)$$

Due to the fact that the frequencies of the test signals are chosen closely together, Eq.(2) can be subsituted in Eq.(1), and thus Eq.(1) can be solved. Equivalent expressions can be derived for signals $\theta^*(\nu)$, $\phi^*(\nu)$ and $\xi^*(\nu)$.
The choice of the best structure of the model is difficult due to the cross-coupling of signals in the simulator system as well as in the human operator. THerefor different structures were used and the results were compared with respect to the quality of the fit of the model to the measured data points. Based on this comparison the following structure is proposed:

$$H_1(\nu)=K_1(1+2\pi j\nu\tau_{11})\,e^{-2\pi j\nu\tau_{v_1}}; \qquad (3)$$

$$H_2(\nu)=K_2\,e^{-2\pi j\nu\tau_{v_2}}; \qquad (4)$$

$$H_3(\nu)=K_3(1+2\pi j\nu\tau_{13})\,e^{-2\pi j\nu\tau_{v_3}}; \qquad (5)$$

$$H_4(\nu)=K_4(1+2\pi j\nu\tau_{14})\,e^{-2\pi j\nu\tau_{v_4}}; \qquad (6)$$

It appears that the complete model has 11 parameters.
To study group effects of alcohol or binary choices on human performances in this set-up, one can consider the averages of the estimated parameters. It is preferable to do the parameter estimation on the average of the raw signals (see also section 5).
An average of two measurements can be taken if the frequencies of the forcing functions in the signals have the same phase in both measurements. In general this is not the case, therefore the different signal components have to be aligned.
This is illustrated in Fig. 7 for the time domain representation.

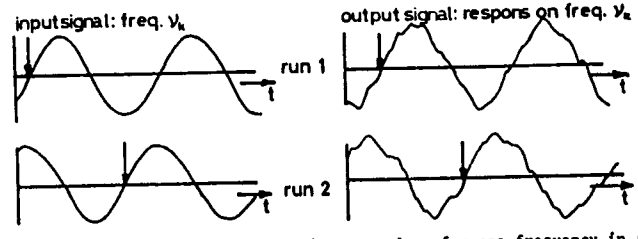


Fig. 7 Illustration of the averaging procedure for one frequency in a signal considered.

In the frequency domain it means a rotation of the vectors representing the Fourier coefficients of the sinusoïdal components.
Various averages are made as will be shown in section 5.

4.2 ANALYSIS OF PHYSIOLOGICAL DATA

From the electrocardiogram the following measures were obtained (Ref. 11).

- Heart rate [beats/minute].
- Standard deviation of the beat to beat duration (RR-intervals) [msec].
- Sum of positive differences of intervals S divided by the number of reversals in increasing to decreasing intervals N: S/N [msec].

These frequently used measures were obtained for a number of periods as shown in Figure 4. All measurements are referred to the first measurement on an experimental day, thus showing a decrease or increase as a percentage of the first value.
The evoked responses could only be recorded just before and after a simulator test run (See Figure 4). The most important parameter in the evoked response is condidered to be the amplitude of the so called top IV. (Ref. 5).

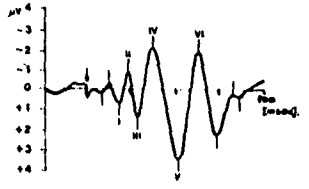


Figure 8. An example of a visual evoked response

Fig. 8 gives an example of an averaged visual evoked response.
The analysis includes the averaging of single evoked responses, smoothing of the average, normalising the individual averages, pooling of the individual averages and calculating the amplitude of top IV.

5 RESULTS

5.1 THE PERFORMANCE DATA

The estimated parameters based on individual measurements of the four test runs and of the six subjects showed a large variance. Due to this fact no significant differences between conditions could be determined.
A better result was gained after averaging the raw signals for the second, third and fourth test run on a day.
This means that three comparable measurements were taken together; i.e. three test runs with a distraction, three test runs with a blood alcohol concentration (BAC) of .05% and three test runs on the reference day of single bicycle task measurements. It is obvious that the averaging of test runs with different BAC and different distraction tasks gives only the possibility to discus general effects of alcohol and distraction tasks.
In Fig. 9 the eleven parameters of the model are given.
The conditions A,B and C are respectively the sessions with distraction task, sessions with alcohol and sessions without capacity reduction.
The means and standard deviations are given for the above mentioned averages of three test runs (solid lines). The figure shows also the estimated parameters (one single value for each parameter) of the average of all comparable test runs (dashed lines).

Table 1 Significant changes in estimated parameters
The levels of significance p are given in the table

H_1	A÷B	A÷C	B÷C
K_1	-	-	-
T_{11}	-	-	-
T_{12}	-	-	-
J_{v1}	p=.10B÷A	-	-
J_1	p=.10B÷A	-	-
H_2	A÷B	A÷C	B÷C
K_2	-	-	-
T_{v2}	-	p=.10C÷A	p=.10 C÷B
J_2	-	-	-
H_4	A÷B	A÷C	B÷C
K_4	-	-	-
T_{14}	-	-	-
T_{v4}	-	p=.10C÷A	-
J_4	p=.05B÷A	p=.10C÷A	p=.05C÷B

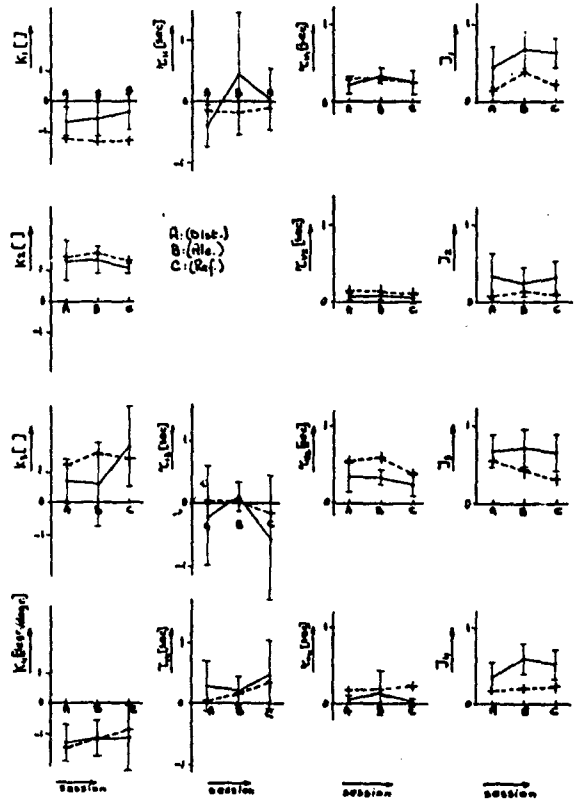


Fig. 9 The parameters of the model of the bicycle rider obtained from averages (of 3 test runs) for each subject separately (solid line) and from the average (of 18 test runs) of all comparable measurements (dashed line).

Furthermore the values of the cost criterion are given, which gives an indication about the quality of the fit to the data points.
Even after averaging it turned out that the cost criterion J_3 remained large thus for $H_3(v)$ no reasonable fit could be made.
The Bode plots and estimated transfer functions of the overall averages are given in fig.10. Finally table 1 shows those differences which turned out to be significant based on a paired t-test. These results will be discussed in section 6.

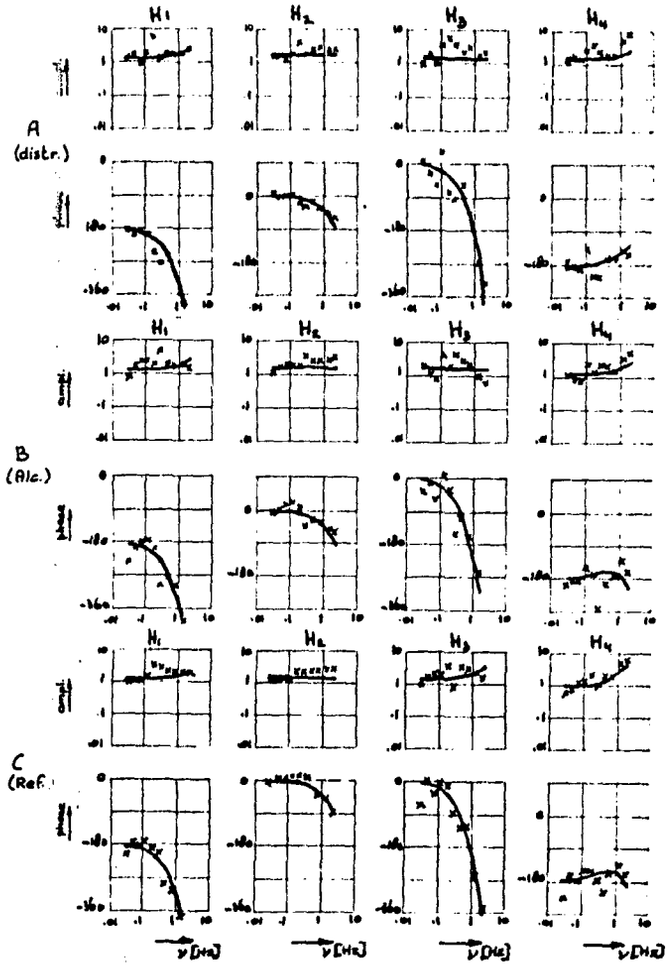


Fig. 10 The Bode plots and estimated transfer functions of the overall average.

5.2 THE PHYSIOLOGICAL DATA

The results of the physiological measurements are summarized in Fig. 11. A very clear difference can be noticed between the values of the ECG

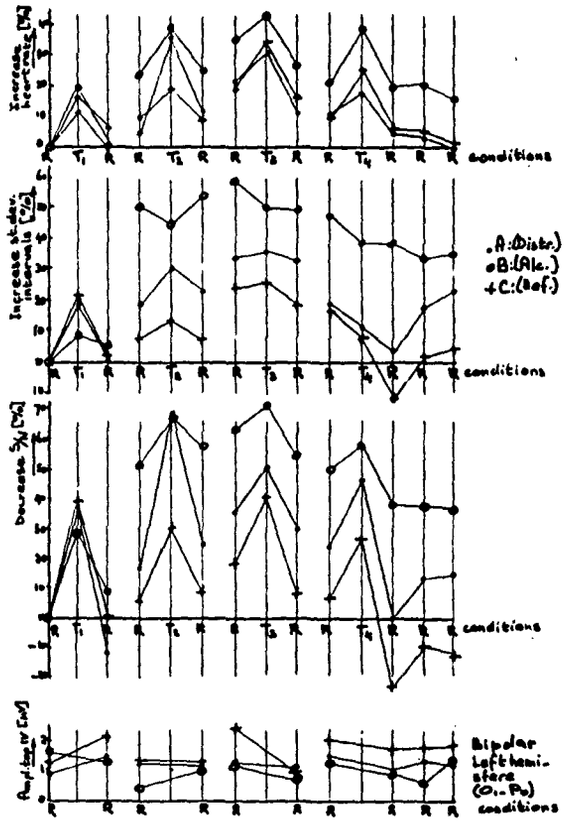


Fig. 11 Increase of heart rate, decrease of heart rate variability and the amplitude changes of top IV in the evoked response during the experimental conditions.

measures between rest and task conditions.
It also seems that the distraction task gives a suppression of the irregularity of heart rate, and the alcohol does even more.
Furthermore a diurnal rhythm is seen in the ECG measures.
The change of top IV in the evoked response does not show clear effects.
The recovery of the amplitude of top IV after the last test run seems to be retarded in the alcohol session.

6 DISCUSSION AND CONCLUSIONS

Two general remarks have to be made about the experimental scheme.

- Each experimental day the same scheme of taking blood samples by venous puncture has been performed. The anticipation of the subjects has probably given a large variance in task performance and also in the physiological measures. This is probably true for interindividual as well as for intra-individual effects.
- There is a lot of experience on parameter estimation in describing function models for one dimensional systems. The situation described in this paper is not known in literature i.e.: a two dimensional system with cross-coupling of the signals in the system to be controlled and the human operator model. Therefore it was impossible to make an optimal experimental design with respect to repeated measures or the number of subjects required.

6.1 THE TASK PERFORMANCE

The subject had the possibility to stabilize the simulator with handle bar corrections or with upper body movements. The course could be followed only with handle bar corrections. It is already mentioned in the introduction that due to interference in the human motor system a coupling between upper body angle and course angle could exist.
However it is concluded that this particular transfer function $H_3(\nu)$ could not be estimated with reasonable reliability.
Based on the results of the transfer functions $H_1(\nu)$, $H_2(\nu)$ and $H_4(\nu)$ the following conclusions can be drawn (see table 1):

- The test runs with distraction task (A) show a smaller time delay in H_1 than those with alcohol (B). The better result of the cost criterion can mean a better course following.
  At the same time the time delay of H_2 and H_4 are larger compared with test runs without capacity reduction (C), thus giving a decrement in the stabilization task.
  The decrease of the cost criterion for H_4 in the runs with distraction indicates that the upper body movements become more important in stabilization.
  A tentative conclusion can be that with the distraction task the course following receives more attention than with alcohol while the stabilization is done with upper body motions based on routine.
  This conclusion is in agreement with the decrease of the time delay in the stabilization with handle bar between distraction runs and the runs without reduction of capacity. Due to a smaller amount of mental capacity in distraction runs it turns out that less mental capacity is given to the stabilization task which can be performed on the level of the spinal cord.
- The time delay constant in H_2 enlarges in the test runs with alcohol (B), with increasing cost criterion for H_4. This seems in contradiction with some literature which states that small blood alcohol concentrations give better performance due to the anticipation of the subjects. An increase of the cost criterion for H_4 in the runs with alcohol indicate that the stabilization is also deteriorated.
  It suggests that the alcohol has a general effect, because of a decrease in stabilization for both aspects i.e.: the control by handle bar and upper body movements. This agrees with the results obtained in earlier studies on the bicycle simulator without the course following task (Ref. 12), there it is suggested that the stabilization is mainly done by the vestibular-cerebellar system. The handle bar motions, particular needed in the coursefollowing, are controlled by higher central nervous system centers, which is suggested to be more flexible and may compensate to a certain degree.
- Large differences between the parameters of individual subjects point to the possibility that each subject has his own strategy.
  Only repeated measurements can give more insight in these individual differences.
- It is seen in Fig. 9 that the value of the cost criterion after averaging the raw signals is much better than the mean value of the cost criterion for the individual estimation.

6.2 THE PHYSIOLOGICAL MEASURES

The main conclusion concerning the physiological measures is that the subjects showed a clear effect in the measures during the test run.
The recovery of the measures after the last test run cannot be interpreted as an effect of taskload, because of the effect of the anticipation on the venous punctures.
It seems better to do experiments without blood sampling during the experimental sessions. Individual blood alcohol curves obtained separately from the experiment can provide enough reliable estimates of blood alcohol concentrations during the experiments.
It is known from literature that alcohol has a general effect on most neurological systems. This agrees with the retarding recovery of the evoked response after the last test run, and also with the tentative conclusions about the performance data.
Finally it has to be stated that the differences in physiological measures of test runs with alcohol and test runs with the distraction task cannot be interpreted, because alcohol has also a pharmacological influence on the myocardial system.

7 ACKNOWLEDGEMENT

The research described in this paper is done in close cooperation with the Man-Machine Systems Group of the Delft University of Technology.
The bicycle simulator is owned by this group. The support of this group, concerning sofware as well as hardware aspects was indispensable in this research.
The EEG investigations were carried out by Dr.J.L.Blom,who had also the responsibility for the medical aspects.

8 REFERENCES

1. ROLFE, J.M.; The measurement of human response in man-vehicle control situation. In: Monitoring behavior and supervisory control. Ed. T.B. Sheridan and G. Johannsen, Nato conference series, Series III: Human Factors, Plenum Press, New York, p.125-137.

2. WELFORD, A.T.; Evidence of a single-channel decision mechanism, limiting performance in several reaction tasks. Quart Journal Exp. Psych., 11, 193-210.

3. KALSBEEK, J.W.H., J.H. ETTEMA; Physiological and Psychological evaluation of distraction stress. Ergonomics,7 (1974) 443-447.

4. KAHNEMAN, D.; Attention and Effort. Englewood Cliffs, N.J., Prentice Hall, (1973) 245 p.

5. BLOM, J.L.; L'influence de la charge mentale sur les potentiels évoqués. Le Travail Humain. 37 (1974) 139-212.

6. ROHMERT, W. et.al.; Heart Rate Variability and Work-Load Measurement. Ergonomics, 16 (1973) 33-44.

7. LUNTEREN, A.van, H.G. STASSEN; Chapter III in Annual Report 1969 of the Man-Machine Systems Group. Delft University of Technology, WTHD 21, (1970) 15-33.

8. KALSBEEK, J.W.H.; On the measurement of deterioration in performance caused by distraction stress. Ergonomics,7 (1964) 339-367.

9. LUNTEREN, A. van, H.G. STASSEN; Parameter estimation in linear models of the human operator in a closed loop with application of deterministic test signals. Proc. of the 9th Annual Conf. on Manual Control, MIT, (1973) 289-297.

10. LUNTEREN, A. van,; Parameter estimation in a human operator describing function model for a two-dimensional tracking task. To be presented at the 13 th Annual Conf. on Manual Control, MIT (1977) 17 p.

11. OMMEER, C.H.J.M.; The information content of successive RR-interval times in the ECG. Preliminary results using factor analysis and frequency analysis. Ergonomics,16 (1973) 105-112.

12. LUNTEREN, A. van, H.G. STASSEN; Chapter IV in Annual Report 1969 of the Man-Machine Systems Group, Delft University of Technology, WTHD 21, (1970) 34-55.

N79-17490

A RELATIONSHIP BETWEEN EYE MOVEMENT PATTERNS AND PERFORMANCE IN A PRECOGNITIVE TRACKING TASK*

D. W. Repperger E. J. Hartzell

Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio 45433

ABSTRACT

A study of eye movements made by various subjects in the performance of a precognitive tracking task is reported. The tracking task presented by an antiaircraft artillery (AAA) simulator has an input forcing function represented by a deterministic aircraft fly-by. The performance of subjects is ranked by two metrics. Good, mediocre, and poor trackers are selected for analysis based on performance during the difficult segment of the tracking task and over replications. Using phase planes to characterize both the eye movement patterns and the displayed error signal, a simple metric is developed to study these patterns. Two characterizations of eye movement strategies are defined and quantified. Using these two types of eye strategies, two conclusions are obtained about good, mediocre and poor trackers. First, the eye tracker who used a fixed strategy will consistently perform better. Secondly, the best fixed strategy is defined as a "Crosshair Fixator".

*The research reported in this paper was sponsored by the Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio 45433. This paper has been identified by the Aerospace Medical Research Laboratory as AMRL-TR-77- . Further reproduction is authorized to satisfy needs of the U.S. Government.

I. Introduction

A problem of interest in the investigation of man-machine interaction is the determination of "good" and "poor" performance when the human is acting as a regulator. Virtually all modeling efforts in this area assume that the operator is acting at his "best" performance level throughout the scenario of operation. A further assumption is that the control strategy is consistent. This requires that the human operator be highly trained and highly motivated in the employment of his regulator strategy. These assumptions may or may not be met, for the human is inherently an information processor and controller. The quality and consistency of his regulator actions may depend on how he initially perceives his information and what type of information he selects for processing. These are operator specific factors contributing to individual differences seen in operator performance. Though impossible to measure the central processing functions of man, it is inviting to attempt to determine how this information processing may manifest itself in the physiological variables amendable to measurement. These would include motor control or hand movements or perhaps the eye movement behavior of the operator.

This approach has motivated the work presented here by considering the variables which could distinguish a "good" from "poor" human regulator. The hand movements certainly are a manifestation of the central processor. However, it is difficult to analyze the manipulator dependent hand response in such a way as to easily classify different types of motions and correlate these with performance. On the other hand, eye movement variables are very close to the central processor and perhaps easier to analyze. If a metric of eye movement patterns can be developed which can be shown to differ significantly with "good" and "poor" operators (in terms of total system performance), then an indication of efficiency or strategy of the human central processor may be obtained. Although this approach is by no means ideal in the sense that it is not directly measuring the central processor, it may yield an indication of the activity at the closest possible level to the brain.

The phase plane was selected as the analytic tool for the development of the eye movement metric. Clark and Stark [1] have recently used the phase plane as a tool in developing a model of saccadic eye movements in a more or less discrete paradigm. In dynamic experiments the phase plane analysis of eye movements correlated with the displayed error signal in a compensatory tracking task enables measures of eye activity to be quantified in an exact and explicit manner. The data presented here represents thoroughly trained subjects in 16 replications of an experimental task which is well known to the subjects.

II. Methods

The Experiment: The task of a gunner in an optically directed antiaircraft artillery (AAA) system provides a relevant opportunity to study the relationships mentioned above. The tracking task may be described as viewing a target aircraft through a magnifying lens system with a restricted field of view (FOV) (see Figure 1). The operator is provided with a cross hair sight reticle in the center of the FOV. The operator's manipulation of a control wheel produces an input signal which drives a plant or controller system which in turn produces the tracking movements of the optical system and gun. The AAA operator is trained to keep the tracking reticle centered on the target aircraft with as little error as the forcing function will permit. The apparent position, velocity, and range of a typical target aircraft relative to the operator are constantly changing as a function of time in a more or less predictable manner. A more thorough description of the experiment can be found in [2] with a preliminary statistical study on the sources and causes of tracking behavior reported in [3].

Subjects: Eight teams of two subjects (both male and female) each were trained on the simulator for approximately three months prior to the experiment. The subjects ranged in age for 18-20 years. The training procedure consisted of 16 replications of the forcing function flyby per day. The subjects had normal or corrected to normal vision and visual field tests showed no abnormalities or limitations.

Task: The tracking task is represented by the combined azimuth and elevation forcing function. The simulated aircraft flys at a fixed attitude of 1524 m and a constant velocity of 232 m past the gun implacement. The simulator flight lasts 50 seconds. The forcing function used in this study resembled the shape of figure (2). The effective plant dynamics in the closed loop controlled by each operator can be represented by the following lumped transfer function:

$$H(s) = \left[\frac{s+1}{s}\right] \frac{64}{s^2 + 12.5s + 64} \qquad (1)$$

Eye Movement Measurement Techniques: Summed horizontal DC electrooculographic (EOG) measurements were made on both the azimuth and elevation operators of each team. The electrodes were secured to the outer canthi of each operator with the indifferent electrode secured to the forehead. The optical system had a 2.3 mm exit pupil and a head rest was available on the binocular system to provide head stability.

III. Segmentation of The Task And Performance Results:

Crank spectrum plots and other metrics indicated that the operators tended to "give up", or regress from tracking during the mid portion (difficult part) of the task owing to the velocity or acceleration profiles of the input forcing function. This observation led to the segmentation of the task into 4 regions for analysis purposes. Figure 3 represents the segmentation of the task and a short description of the phases of tracking behavior.

The input forcing function velocity profile indicated that the azimuth task is about 4 times more difficult than the elevation task. Thus, only the azimuth task is reported here. Mean square tracking errors, ensemble mean error, spectrum analysis of error, and crank power suggested that the tracking error and eye movement patterns of 3 of the teams were representative of the entire subject complement and were selected for more extensive analysis. The tracking performance hierarchy supported by all metrics showed team 7 the best, with teams 5 and 2 following in descending order. Table 1 summarizes these results for both the e_{RMS}

performance metric and the $\sigma_{\text{Hit Score}}$ metric. $\sigma_{\text{Hit Score}}$ was defined as the amount of time the error signal was within a specified window width.

Table 1

Performance Hierachy of Azimuth Operators

	σ_{RMS}	$\sigma_{Hit\ Score}$
PHASE 1	7 5 2	7 mixed
PHASE 2	7 5 2	7 5 2
PHASE 3	7 5 2	7 5 2
PHASE 4	7 5 2	mixed

IV. Time Series of Eye Movements:

Figure (4) illustrates the real time series plots of the signals e(t) and the EOG signal of team (5) (the mediocre tracker). One observes from these ensemble mean plots that the mediocre tracker has eye movements which follow the stage error position displayed to him. This is especially true during the difficult segment of the tracking task (18-35 seconds). We classify this type of eye tracking strategy as "Eye Follower". In this case the eye follows the stage position.

Figure (5) illustrates the real time series plots of the best tracker in this study. It is observed that the eye movements remain essentially around zero even under the condition of large values of the error signal. This type of eye strategy is defined as "Cross Hair Fixator". In this case when the target becomes difficult to follow (18-35 seconds), the tracker fixates on the cross hair and views the stage in the periphery of the eye.

It is noted that a cause effect question can be raised as to whether the good tracker's eye movements (near zero) may be due to the fact that his error signal is small resulting in small eye movements. Although this is a good point, it is not the case as can be seen in figures (4) and (5). During the difficult segment of the tracking task (18-35 seconds), the good tracker had an error signal which was approximately 80% to 90% that of the mediocre tracker. In other words, the difference in performance of these two trackers was not substantial. The eye movement patterns, however, differed by a factor of 5 to 1 as can be seen in figures (4) and (5).

Figure (6) illustrates the time series of eye movements and tracking error for the poorest tracker in this study. This type of eye tracking strategy is inconsistent and mixed. During the less difficult segments of the tracking task this operator uses his eyes to follow the target. During the difficult segment of tracking, the eyes change strategy and they do not follow the stage position. We define this type of eye tracking as "Inconsistent".

It is noted that all trackers achieved these eye movement patterns after an extensive training period of over 2.5 months. Of all the subjects involved in this experiment, the team 5 tracker had two and one half years experience in similar manual control tasks prior to the 2.5 months training for this study. Recall, however, that the team 5 operator was only a mediocre tracker.

V. A Phase Plane Analysis of The Error and EOG Signals:

Figure (7) illustrates the phase plane plots of the time error signal e(t) and the EOG signal of the mediocre tracker. This figure can be compared to figure (4) in real time. One observes that the "Eye Follower" has a phase plane pattern of EOG movements very similar in shape to the same phase planes of the error signal.

Figure (8) illustrates the phase plane plots of the same type as in the previous figure but for the good tracker. This plot may be compared to the real time series in figure (5). It is noted that the "Crosshair Fixator" has an eye

movement phase phase plane with erratic behavior around the origin. This result is consistent with the time series analysis.

Figure (9) is the same phase planes for the poor tracker. In this case the mixed or inconsistent strategy is observed to occur. The eye movement phase plane in figure (9) appears to be a combination of eyemovements of figures (7) and (8) and demonstrates the inconsistency in eye movement patterns.

Since the results presented here so far have been based on heuristic conclusions, it is desired to quantify these patterns in some manner so that more explicit conclusions can be stated.

VI. A Methodology of Quantification of Eye Movement Strategies:

With reference to the phase plane figures (7,8,9), it is desired to quantify the similarities and differences between eye movements and the error signal. Figure (10) illustrates a vector measure in the phase plane considered here.

Taking the ratio of the vectors provides a metric of changes in the phase planes. This metric suggests that a ratio can be established to differentiate between operators who follow the error or fixate on the cross hair. The ratio suggested is:

$$\bar{R} = \frac{R_{EOG\ eye\ movement}}{R_{ERR\ displayed\ error}} \qquad (2)$$

For the case of the cross hair fixator we would expect:

$$\bar{R} \approx 0 \qquad (3)$$

for the change in eye movement which is very small compared to the change in the error signal. While in the case of the eye follower, the ratio should satisfy:

$$\bar{R} \approx 1 \qquad (4)$$

Figure (11) represents the mean and standard deviations at the 4 time periods of the phase plane vector ratios of team 5 (the mediocre tracker). From the phase plane we would expect that the condition $\bar{R} \approx 1$ should be satisfied. Using the mean value of $\bar{R}$ we see that this is the case and the eye movement and error signal are highly correlated.

For the case of team 7, we would expect $\bar{R} \approx 0$ for the crosshair fixator. From figure (11) we see this is the case from the mean values of $\bar{R}$. Also it should be noted from figure (11) the decay of the standard deviations in phases 2 and 3 which is typical of all teams. The team 7 operator who tends to fixate the cross hair and has, at the same time, the lowest RMS error, presents a further complication. There is no significant difference shown by "t" tests between the RMS error and zero in segments 1 and 4 over the 16 replications of the task. Further, "t" tests show no significant difference from zero in the eye movement signal. Thus, the vector ratio is composed of two very small numbers which merely represent noise and indeed the mean ratio in segments 1 and 4 approximate 0.5. In segments 2 and 3, $\bar{R} < 0.2$. In phases 2 and 3, however, the differences between the good and poor tracker is more sharply defined. Since this is the region where the task is most difficult, this would be the period of most interest to study the eye movements.

VII. Conclusions:

The results of this study indicate two important points. First, the trackers with the consistent strategies (Eye Followers and Crosshair Fixators) had better performance than the tracker who was inconsistent with his eye movements. The second important point is that of the two consistent strategies, the best tracker is a Crosshair Fixator.

It is also noted that the above results are independent of the extent of training. The tracker who was classified as the Eye Follower had two and one half years of experience with motor control tasks. In addition, he had almost 3 months of training on this task, tracking 16 flybys on a daily basis.

References

[1] Clark, M. R. and L. Stark, "Sensitivity of Control Parameters in a Model of Saccadic Eye Tracking and Estimation of Resultant Nervous Activity", Bull. of Mathematical Biology, Vol. 38-1, pp. 39-57, 1976.

[2] Repperger, D.W. and E.J. Hartzell, "Performance Measures of Human Tracking Utilizing PID Modeling and a Closed Loop Error Metric", NAECON, pp. 887-893, 1976.

[3] Repperger, D.W., W.C. Summers, E.J. Hartzell, and G.D. Callin, "A Phase Plane Approach To Study The Adaptive Nature of A Human Performing a Tracking Task", 1976 IEEE Conference on Decision and Control, Clearwater, Florida, December, 1976.

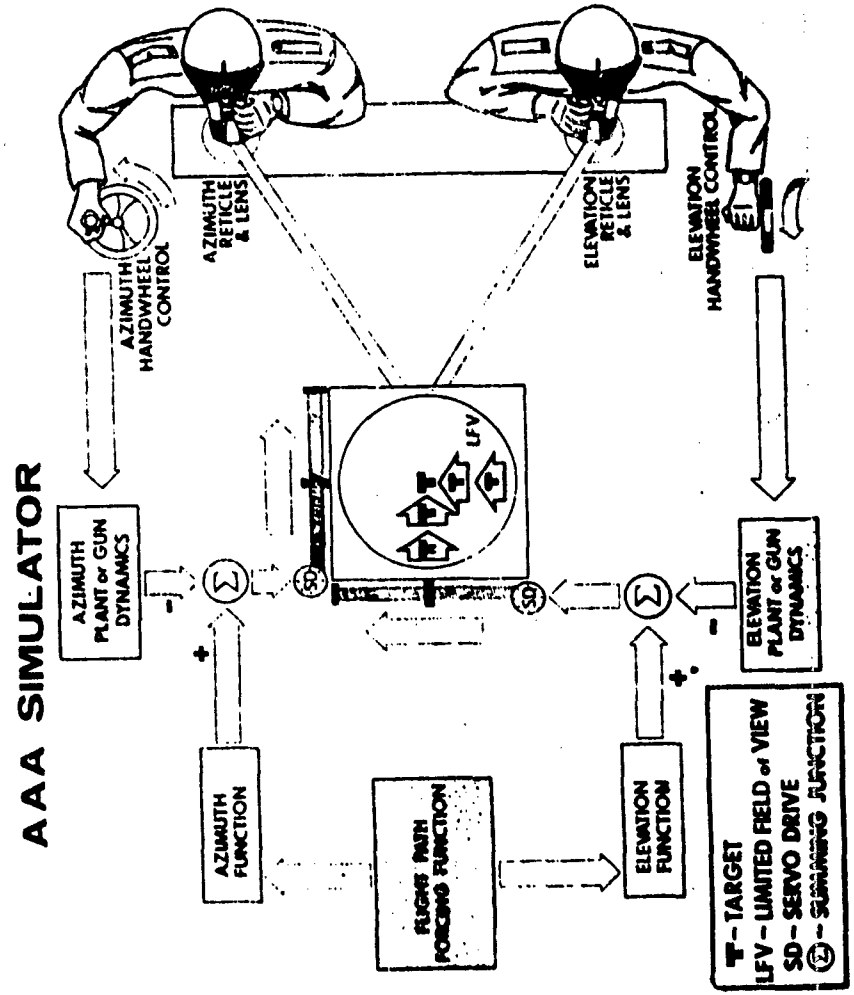


Figure (1)

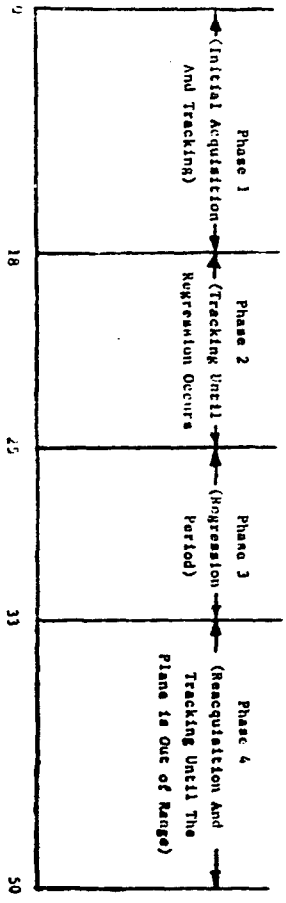


Figure (3) The 4 Phases of Tracking The Flyby Trajectory

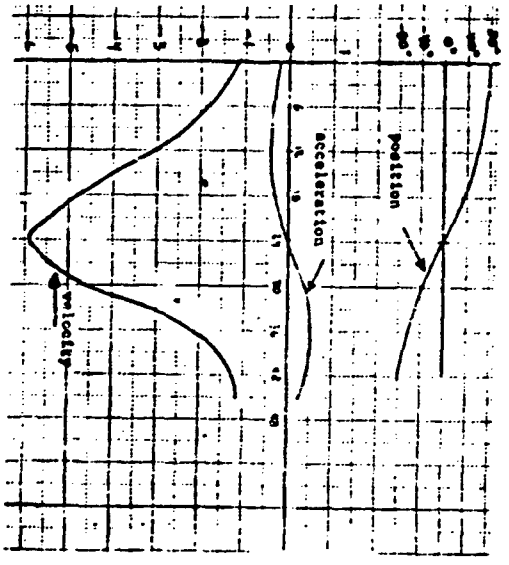


Figure (2)

The Forcing Function Input

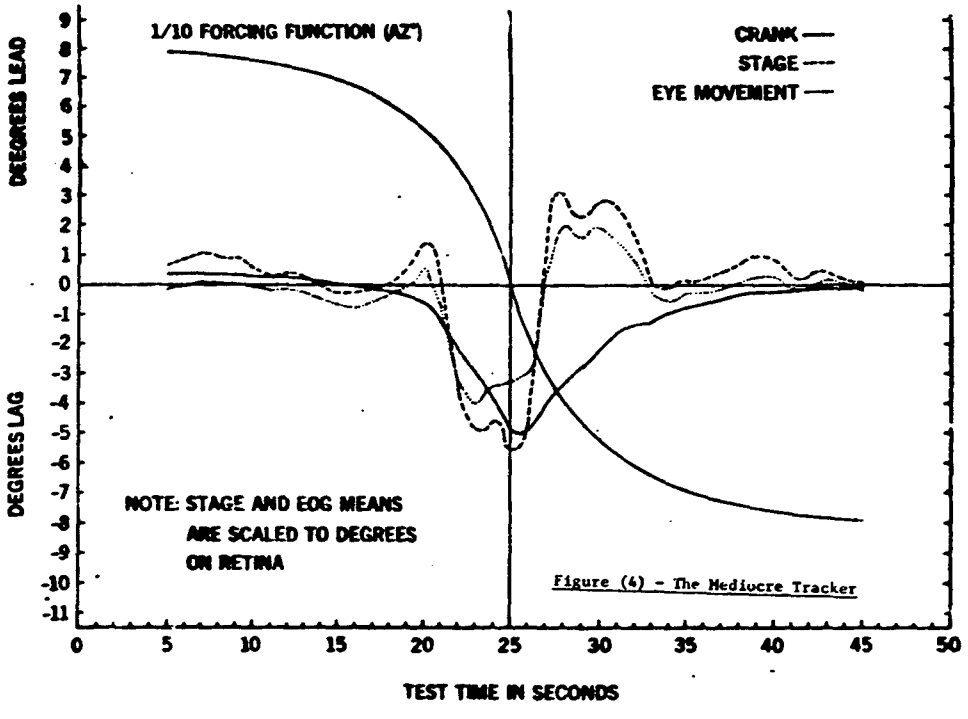
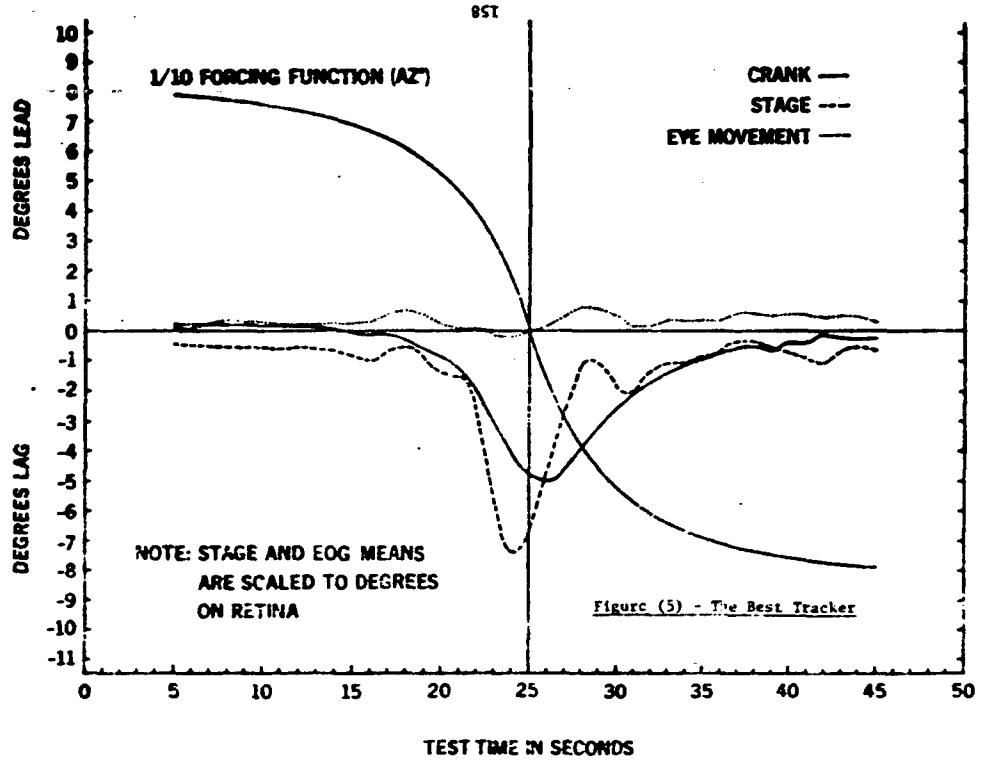
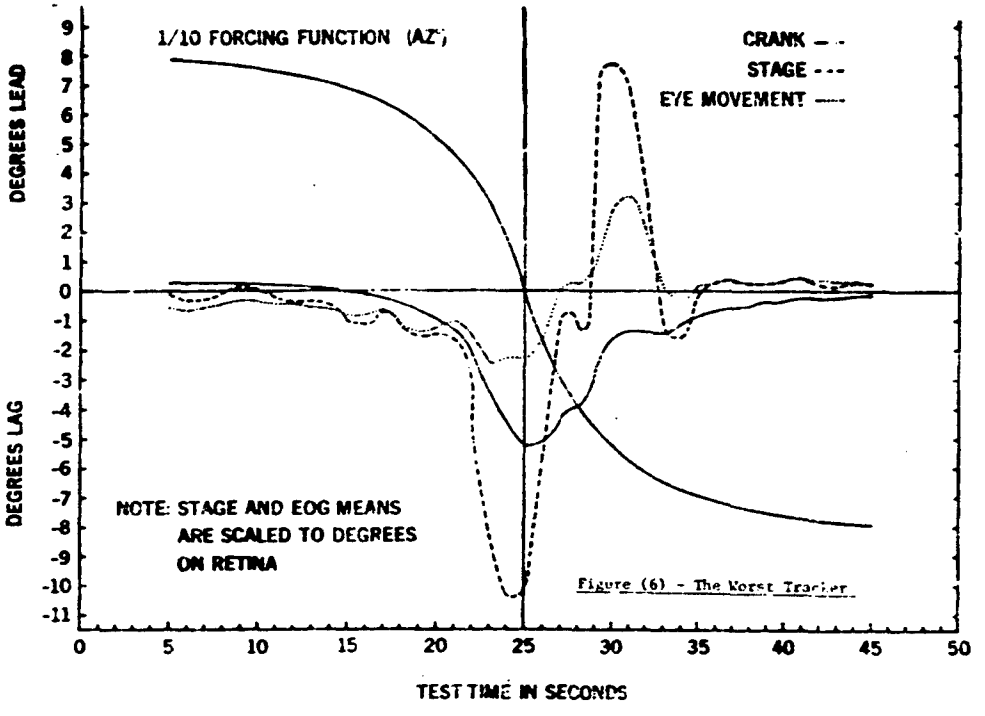


Figure (4) - The Mediocre Tracker

TEAM 5 AZ OPERATOR MEAN ERROR DATA
(STAGE-CRANK-EYE MOVEMENT EOG)



TEAM 7 AZ OPERATOR MEAN ERROR DATA



TEAM 2 AZ OPERATOR MEAN ERROR DATA
(STAGE-CRANK-EYE MOVEMENT EOG)

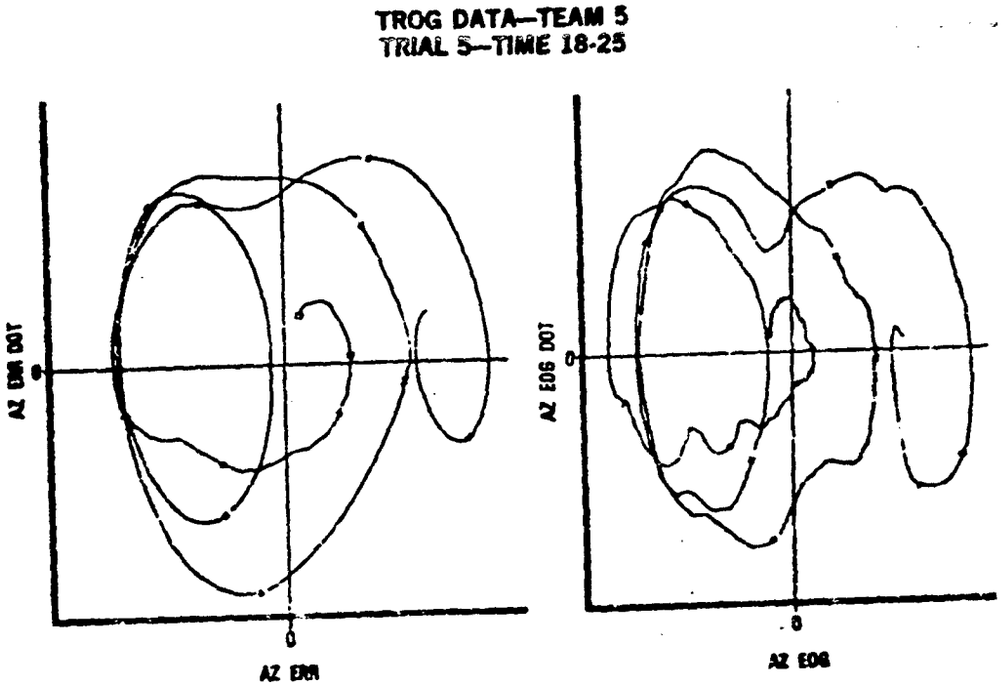


Figure (7) - The Mediocre Tracker (Eye Follower)

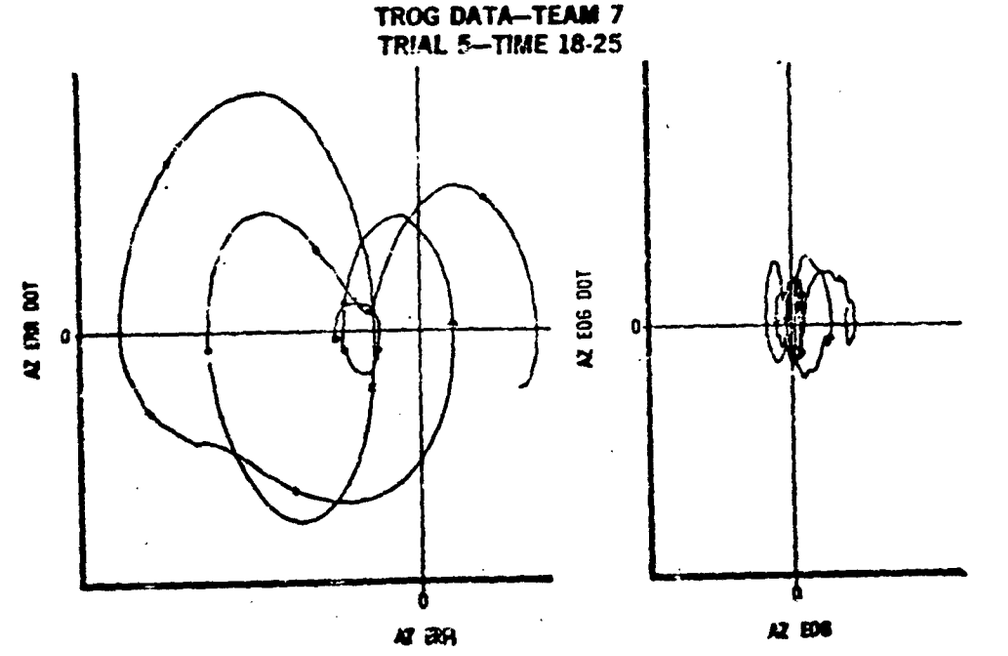


Figure (8) - The Best Tracker (Eye Fixator)

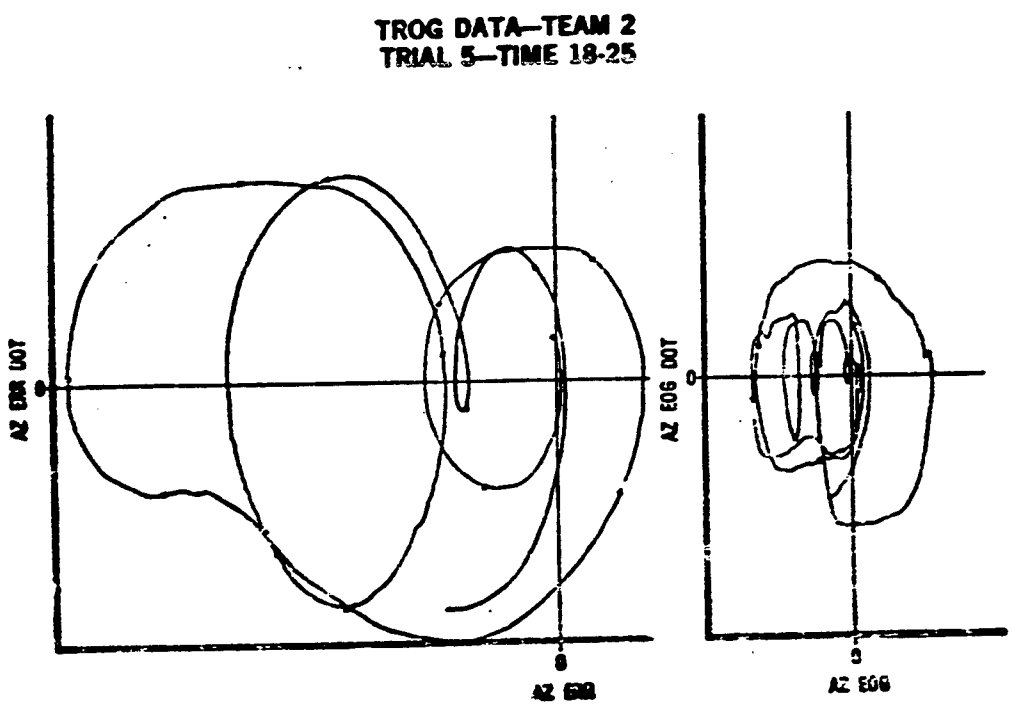


Figure (9) - The Worst Tracker (Inconsistent Eye Strategy)

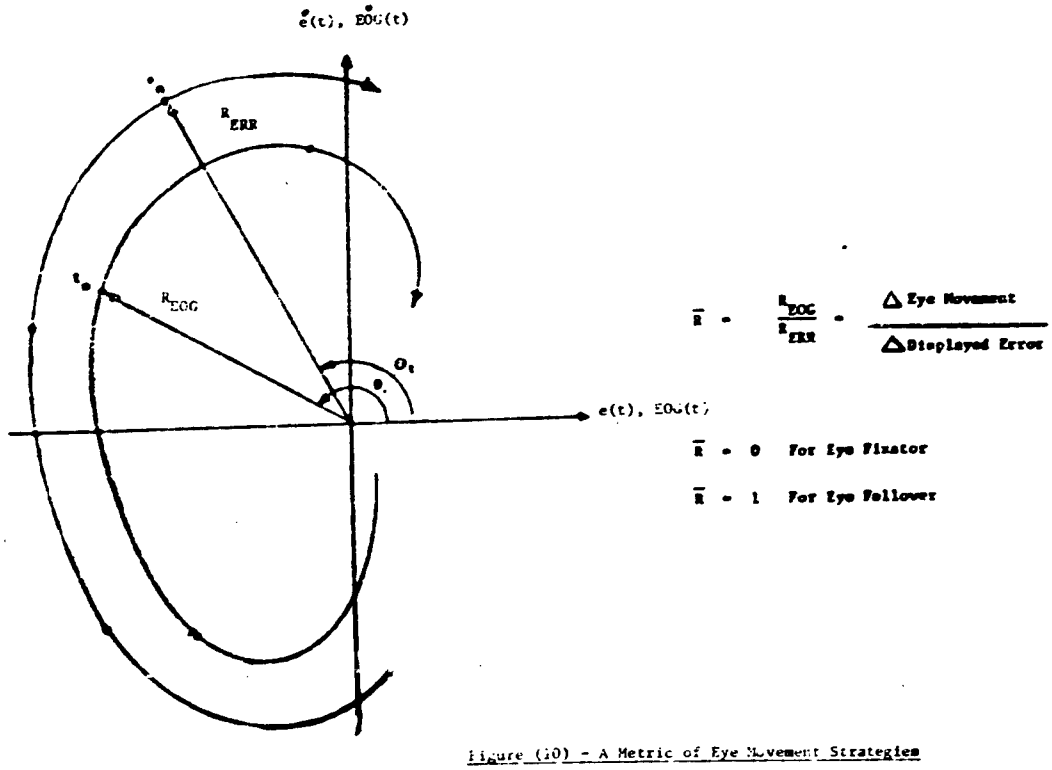


Figure (10) - A Metric of Eye Movement Strategies

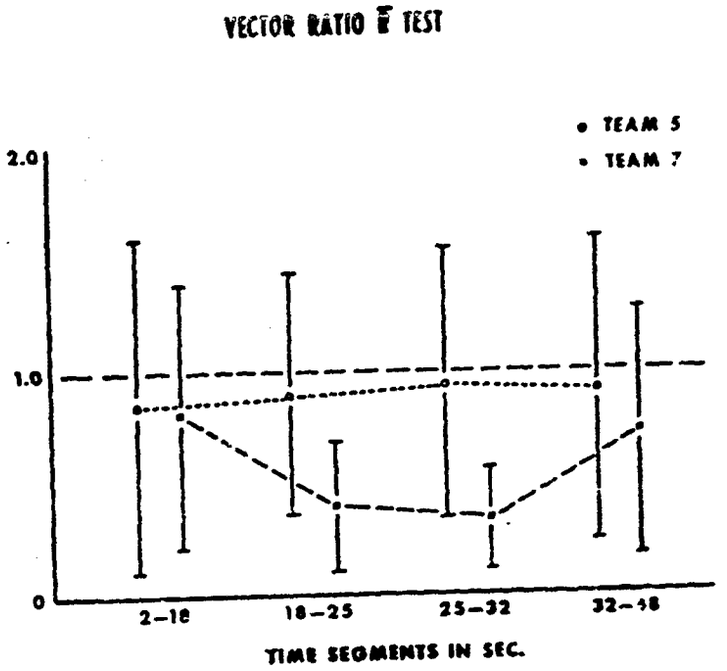

VECTOR RATIO $\bar{E}$ TEST
• TEAM 5
• TEAM 7
2.0
1.0
0
2–18
18–25
25–32
32–48
TIME SEGMENTS IN SEC.

Figure (11)

Session III

SURFACE VEHICLE CONTROL

Chairman: T. B. Sheridan

A CONTROL THEORETIC MODEL OF DRIVER STEERING BEHAVIOR

Edmund Donges

Forschungsinstitut für Anthropotechnik
Meckenheim, Germany

Abstract

In establishing design criteria for vehicle dynamics which may improve the performance of the driver-vehicle system, a quantitative description of driver steering behavior such as a mathematical model is likely to be helpful.

The steering task can be divided into two levels: (1) the guidance level involving the perception of the instantaneous and future course of the forcing function provided by the forward view of the road, and the response to it in an anticipatory open-loop control mode; (2) the stabilization level whereby any occuring deviations from the forcing function are compensated for in a closed-loop control mode.

This concept of the duality of the driver's steering activity led to a newly developed two-level model of driver steering behavior. Its parameters were identified on the basis of data measured in driving simulator experiments. The parameter estimates of both levels of the model show significant dependence on the experimental situation which can be characterized by variables such as vehicle speed and desired path curvature.

1. Introduction and Conception of a Driver Model

The long term objective of investigating dynamic characteristics of the driver-vehicle system is to establish criteria for the design of vehicle dynamics which may improve the "Active safety" of vehicles. To accomplish this the dynamics of both system elements, i.e., driver and vehicle, have to be known. While the mathematical description of vehicle dynamics is highly developed, there is a considerable lack of such knowledge about the dynamic capabilities and limitations of the driver. In order to help improve the description of driver behavior, the steering activity of drivers measured in an extensive series of driving simulator experiments was investigated. The results are presented in the comprehensive form of a mathematical model.

The basic idea underlying the data analysis and the driver-model concept is the duality of information presented to the driver by the forward view of the road. On the one hand, the visual field of the driver provides information on the instantaneous and future course of the road, so that the driver can extrapolate not only the present but also the future run of the driver-vehicle system's forcing function. This type of visual perception is called anticipation or preview [7], [4], [10], [12]. The information on the forcing function is called "Guidance information", because the driver uses it for guiding the vehicle along its desired path. A physical quantity which seems suitable for the description of the sensory impression about the forcing function, is the curvature of the road (i.e. the reciprocal of the radius of turn), because the central perspective pattern of the road shows a straight or a more or less curved run corresponding to the value of curvature. Therefore the curvature of the vehicle's desired path can be looked at as a measure of the forcing function.

On the other hand, static and dynamic cues in the visual field contain information on the instantaneous deviations between the vehicle's actual path and its desired path (i.e, forcing function). This portion of information is called "Stabilization information", because the driver uses the corresponding visual cues to stabilize vehicle motions with respect to the forcing function. The essential portion of stabilization information can be described by the following quantities (see figure 1):

1. Lateral deviation y_Δ between the driver's head position and the vehicle's desired path.
2. Angle between the tangents of the desired and the actual path. Because in the simulated vehicle dynamics the tangent of the vehicle's actual path is identical with its longitudinal axis, this quantity is called heading angle error ψ_Δ.
3. Difference of the curvatures of the vehicle's actual and desired path, called path curvature error K_Δ.

The way these three quantities may be perceived by the driver, is discussed in [3] and [4] with reference to [1], [6] and [7].

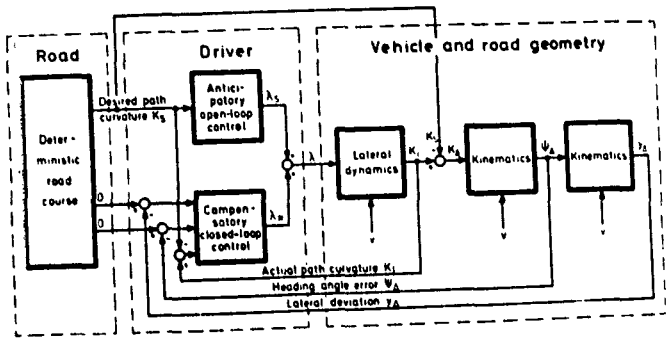


Fig. 1. A structural scheme of human steering behavior in the driver-vehicle-road system

The duality of the driver's visual information is the basis of the two-level structure of the driver model described in this paper (Figure 1). Accordingly, the two levels of the model are called "Guidance level" and "Stabilization level". Instead of a sequential mode of operation as proposed by Crossman and Szostak [2], both levels operate here in parallel. At the guidance level, the anticipation of the forcing function enables the driver to steer in an anticipatory manner when the forcing function changes. Because an anticipatory reaction is always feedforward, such steering behavior can best be explained as an "Anticipatory open-loop control". At the stabilization level, the success of this open-loop control is observed and must be completed by compensation for any occuring deviations. This "Compensatory closed-loop control" stabilizes vehicle motion.

2. Driving Simulator and Experimental Procedure

The experiments providing the data for this investigation were carried out in the driving simulator of the Forschungsinstitut fuer Anthropotechnik (FAT) in Meckenheim, Germany [9]. This driving simulator consisted of an analogue computer simulation of longitudinal and lateral vehicle dynamics, a mock-up of the driver's seat with speedometer and control elements (steering wheel, gas and brake pedals, automatic gearshift), and special subsimulators for the generation of driving and engine noise, vehicle motions and the visual out-of-the-window scene (Figure 2).

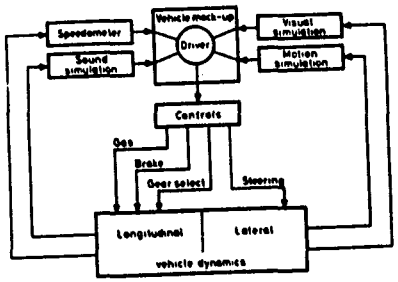


Fig. 2. Topology of the FAT driving simulator

The simulated test course was a winding closed-circuit two-lane road located in a horizontal plane (Fig. 3). The length of the course was 3.2 km. The route consisted of a series of joined road sections of constant curvatures (Figure 3 b).

The visual scene which contained the essential information for the driver was a simulated quasi-natural out-of-the-window forward view of the two-lane test course (Figure 4). This road image was generated according to the instantaneous position and attitude of the vehicle by electronic means and shown to the driver on a 3 m high and 4 m wide screen by a black and white TV projection. The horizontal width of the visual field was about 40°. The angular relations of the lane markings in the simulated visual field as seen by the driver corresponded to an 8 m wide road.

The analogue computer representation of the vehicle dynamics used for this investigation approximately simulated an automobile. The lateral dynamics

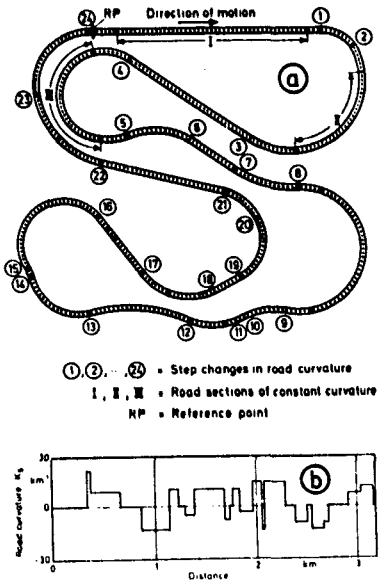


Fig. 3. Test course
a) Plan view
b) Curvature profile of road centerline

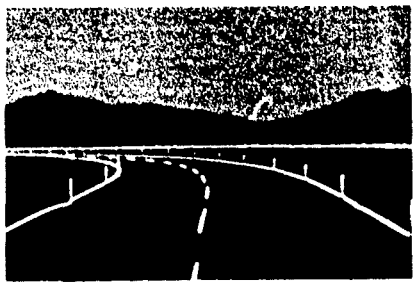


Fig. 4. Quasi-natural visual simulation of a two-lane road

of the vehicle can be defined by the relationship between the steering wheel angle λ and the vehicle's actual path curvature K_I [5]. In the ranges of lateral acceleration and of steering wheel rotation frequencies covered during the experimental runs, the vehicle's lateral dynamics can be characterized *in a first order approximation* by a time-delay term:

$$K_I(t) = E_L \, \lambda(t-\tau_F) \quad . \tag{1}$$

Herein, $\tau_F = 0.2$ s is the vehicle's time-delay constant (vehicle-response time) and E_L is the steady state steering sensitivity ($E_L = 0.85 \cdot 10^{-3}$ degree^{-1}m^{-1} = 0.049 rad^{-1}m^{-1}). Since the vehicle is assumed to be neutral steering, E_L is independent of vehicle speed [8]. It should be mentioned that the value of E_L used here is about two to three times as high as usually applied in automobiles.

When side slip angle is neglected, the following kinematic equation describes heading angle error ψ_Δ, i.e., the angular difference between the desired path's tangent and the vehicle's longitudinal axis:

$$\psi_\Delta(t) = \psi_\Delta(t_o) + \int_{t_o}^{t} v(\vartheta) \left[K_I(\vartheta) - K_S(\vartheta)\right] d\vartheta \tag{2}$$

A corresponding equation exists for the kinematic relationship between heading angle error ψ_Δ and lateral deviation from the vehicle's desired path y_Δ:

$$y_\Delta(t) = y_\Delta(t_o) + \int_{t_o}^{t} v(\vartheta) \, \psi_\Delta(\vartheta) \, d\vartheta \quad . \tag{3}$$

In equations (2) and (3), $\psi_\Delta(t_o)$ and $y_\Delta(t_o)$ are initial conditions, v is the vehicle's speed and K_S is the curvature of the vehicle's desired path.

To have a well defined, measureable quantity K_S as the forcing function of the driver-vehicle system, all subjects were insistently instructed to drive along the centerline of the road as accurately as possible, i.e., while driving, the subjects should make sure that they sit just above the road centerline. During the experiments the simulated road was free of other traffic.

The data analysis and model identification described in the following chapters is based on data from six subjects measured during experimental runs of three laps on the test course of figure 3.

3. Proof of the Driver's Anticipatory Steering Reaction

As shown in figure 3.b, the test course consisted of 24 sections of constant curvatures. At points where two of these sections join together, the road curvature abruptly takes a new value. This change in road geometry is seen by the driver a long time before he passes the joining point, so that he is able to react beforehand. Typical time histories of such steering reaction and of the corresponding output variables are shown in figure 5. The 4th step change of road curvature (see figure 3) is taken as an example. The diagrams show means and standard deviations averaged over 18 individual time histories.

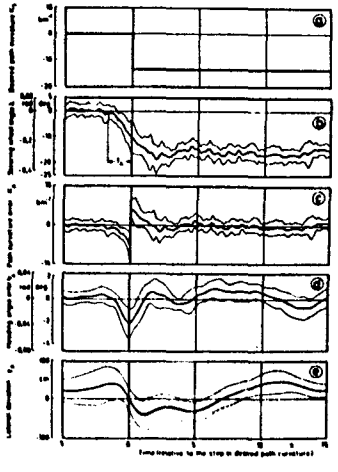


Fig. 5. Typical time histories of system variables at the 4th step change in desired path curvature
a. desired path curvature
b. steering wheel angle
c. path curvature error
d. heading angle error
e. lateral deviation

The steering wheel reaction starts at a certain time interval, called "anticipation time" T_A, prior to the step change of road curvature and then shows a lagging transient with a small overshoot (Figure 5.b). The success of this reaction is most clearly demonstrated in the graph of lateral deviation (Figure 5.e): the change in lateral deviation caused by the step change of road curvature is distributed approximately evenly between right and left hand sides of the forcing function.

4. Driver Model Structure

The complete structure of the two-level model is shown in figure 6.

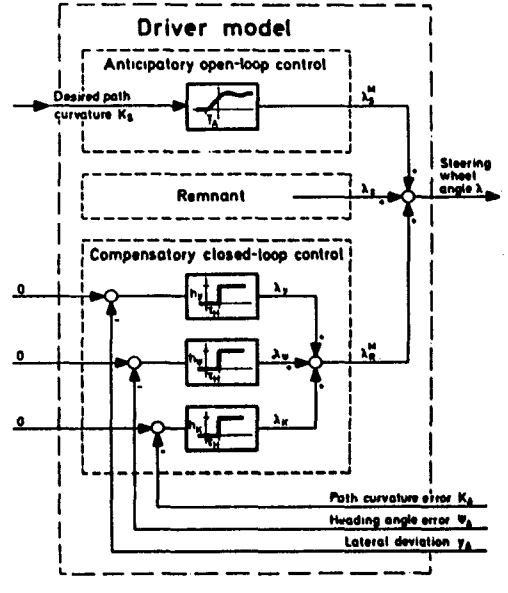


Fig. 6. Block diagram of the two-level model of driver steering behavior

The "Anticipatory open-loop control" submodel representing the guidance level of steering activity simulates the anticipatory response to the deterministic run of desired path curvature. Corresponding to the typical step response of steering wheel angle in figure 5.b, the features of this submodel include anticipation time T_A and lagged transient response. Because the experimental data were registered in sampled form, the anticipatory submodel is formulated in the discrete time form of a scalar difference equation:

$$\lambda_S^M(t_k) + a_n \lambda_S^M(t_{k-1}) + \ldots\ldots + a_1 \lambda_S^M(t_{k-n}) = b_n K_S (t_{k-1} + T_A) + \ldots\ldots + b_1 K_S (t_{k-n} + T_A) \quad (4)$$

where λ_S^M is the steering wheel angle of the anticipatory submodel, K_S is the desired path curvature, t_k ($k = 1, 2, 3, \ldots$) are the discrete points in time, T_A is the anticipation time, a_i, b_i ($i = 1, 2, \ldots, n$) are the time-invariant, linear parameters of the difference equation, and n is the order of the difference equation.

The "Compensatory closed-loop control" submodel corresponding to the stabilization level of steering activity accomplishes the feedback of the output variables. For the mathematical formulation of the compensatory submodel, a very simple approach is used. Each of the three output quantities, i.e., path curvature error K_Δ, heading angle error ψ_Δ, and lateral deviation y_Δ is delayed by an inherent human time delay and fed back by a single gain factor:

$$\lambda_R^M(t) = -\left[h_K K_\Delta(t-\tau_H) + h_\psi \psi_\Delta(t-\tau_H) + h_y y_\Delta(t-\tau_h)\right] \quad (5)$$

where λ_R^M is the steering wheel angle of the compensatory submodel, h_K, h_ψ, and h_y are the gain factors and τ_H is the human time delay.

It has been proved that this simple approach provides the capability of stabilizing the outer loop, i.e., the lateral deviation loop, of the compensatory control process which then shows the features of the crossover model of the human controller [11].

Because both submodels will not completely reproduce the driver's steering wheel angle, a remnant quantity is a third integral part of the model. The remnant comprises driver-induced signals which are not related to the steering task as well as shortcomings of the special structure and features of the model chosen.

At 13 specific sections of the closed-circuit test course, the parameters of the driver model are identified from the experimental data using measurement intervals of 20 s duration [4]. Within this time, vehicle speed is approximately constant, so that a model with piecewise time-invariant parameters is applicable. Ten of the sections contain at least one step change of road curvature which can be utilized for the identification of both model levels. Three sections (I, II, and III in figure 3) have constant curvatures where only compensatory submodel parameters can be identified. Within each section, the experimental data of six subjects with three laps each, i.e., 18 sets of data per section, were evaluated.

5. Results and Discussion

In order to provide a qualitative impression of model performance, the steering angles of the driver and the driver model are compared, and the components of the model steering angle shown by way of an example (Figure 7).

Figure 7.a indicates the experimental conditions. The road section under consideration shows three step changes in road curvature, the vehicle speed is approximately constant. As the comparison of driver and driver model steering angles in figure 7.b shows, the model simulates the driver's steering reactions very well especially as far as the low frequency components are concerned. Figure 7.c presents the contributions of the anticipatory and the compensatory submodels. In figures 7.d to f, the compensatory submodel components are shown. The model remnant in figure 7.g turns out to be a broad band stochastic signal.

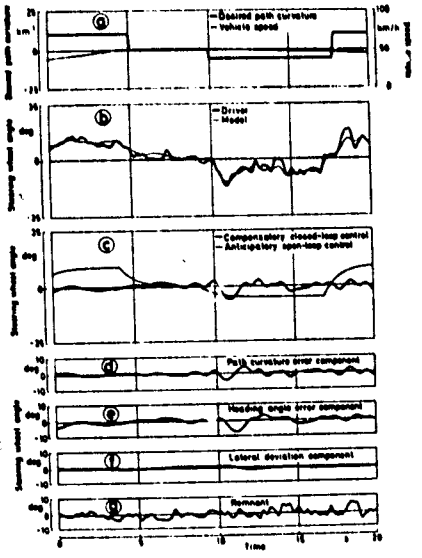
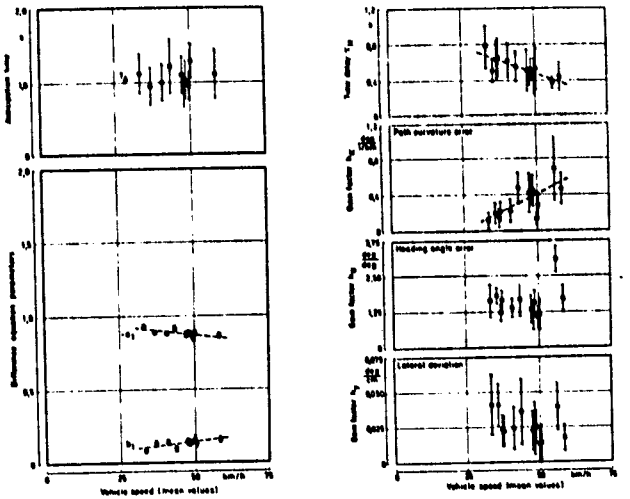


Fig. 7. Time histories of steering wheel angles of driver, driver model, and driver-model components as illustrated by the example of subject 09, lap 3, area of road curvature step changes Nos. 6, 7 and 8 (Figure 3)

Within each of the 13 road sections mentioned above, the experimental conditions for all subjects and all laps are equal, but they differ between the different sections. These differences can be described by experimental variables such as road curvature and vehicle speed. Various parameters of both model levels are significantly related to those experimental variables.

As an example, model parameter estimates are shown as functions of vehicle speed in figure 8. The mean values of 18 individual estimates per section for each parameter are plotted with their 99 %-confidence intervals. The significant relationships are indicated by the linear approximations of the dotted regression lines.



a. b.

Fig. 8. Model parameters as functions of vehicle speed
Mean values with 99 %-confidence intervals
a. "Anticipatory open-loop control"
Order of difference equation: n = 1
b. "Compensatory closed-loop control"

Changes in the order of the difference equation (Eq. (4)) between n = 1 and n = 5 turned out to have little influence on "Anticipatory open-loop control" submodel performance [4]. Therefore, the parameters presented in figure 8.a are for the simplest case of order n = 1. While a definite influence of vehicle speed with respect to anticipation time could not be discovered, both linear coefficients a_1 and b_1 of the difference equation are significantly related to vehicle speed. The evident consequence of these relationships is an increase in the initial slope of the difference equation's step response with increasing vehicle speed.

Two of the parameters of the "Compensatory closed-loop control" submodel also show significant dependencies with respect to vehicle speed (Figure 8.b). The model's time delay decreases as vehicle speed increases. This relationship is reasonable, because the driver has spare time while driving at low speeds as opposed to higher speeds when he is forced to react faster. This adaptation in

time delay is constrained by some lower reaction time limit. Therefore, the relationship indicated by the dotted regression line cannot be extrapolated as in the form of a linear law beyond the speed range covered in the underlying experiments.

The second relationship of the compensatory submodel parameters as illustrated in figure 8.b by path curvature error gain also seems to be reasonable : The velocity vector field [6], [7] of the driver's visual field, which is the basis for the perception of path curvature error [4], becomes more observable as vehicle speed grows. The driver is then able to feed path curvature error back with higher gain.

In addition to the tendencies mentioned which demonstrate the high adaptability of the driver, figure 8.b shows an item which may indicate a human limitation. In one of the road sections, the gain factor of heading angle error is significantly higher than in all other sections. This outstanding section is the only straight road section (Section I in figure 3). The perception of heading angle error in curved road sections is degraded by the fact that the instantaneous direction of the desired path is not directly depicted in the visual field as it is on straight roads [3], [4]. This degradation causes a decrease in heading angle error gain which is disadvantageous from a control theoretic point of view.

6. Conclusions

The results provided by the two-level model of driver steering behavior indicate human adaptability as well as human limitations. The experimental variables in the present investigation which caused variations in human control behavior were vehicle speed and road curvature.

For the initially mentioned long-term objective of improving driver-vehicle-road system performance, the influence of vehicle dynamic parameters such as position of the center of gravity, or tyre features, on human steering activity is of great interest. The two-level model and its identification procedure seem to be suitable for (1) describing human steering behavior, and (2) theoretically evaluating the dynamic interaction within the driver-vehicle control system.

7. References

[1] Biggs, N.L. : Directional Guidance of Motor Vehicles - A Preliminary Survey and Analysis. Ergonomics 9 (1966), pp. 193 - 202.

[2] Crossman, E.R.F.W., and Szostak, H. : Man-Machine Models for Car Steering. In : Proc. 4th Annual Conf. on Manual Control, NASA SP-192 (1968), pp. 171 - 195.

[3] Donges, E. : Experimentelle Untersuchung des menschlichen Lenkverhaltens bei simulierter Straßenfahrt. Automobiltechnische Zeitschrift, 77 (1975), pp. 141 - 146, 185 - 190.

[4] Donges, E. : Experimentelle Untersuchung und regelungstechnische Modellierung des Lenkverhaltens von Kraftfahrern bei simulierter Straßenfahrt. Dissertation Technische Hochschule Darmstadt 1977.

[5] Fiala, E. : Zum Lenkverhalten von Kraftfahrzeugen. Automobiltechnische Zeitschrift, 72 (1970), pp. 111 - 116.

[6] Gibson, J.J. : The Perception of the Visual World. Houghton Mifflin Co., Boston, 1950.

[7] Gordon, D.A. : Perceptual Basis of Vehicular Guidance. Public Roads 34 (1966), pp. 53 - 68.

[8] Hoffmann, E.R., and Joubert, P.N. : The Effect of Changes in Some Vehicle Handling Variables on Driver Steering Performance. Human Factors 8 (1966), pp. 245 - 263.

[9] Schulz-Helbach, K.D., and Donges, E. : On Steering Dynamics of Tracked Vehicles - Results of an Anthropotechnical Investigation by using a Novel Driving Simulator. In : Proc. Symposium on Psychological Aspects of Driver Behaviour, Noordwijkerhout, Netherlands, August 2-6, 1971.

[10] Sheridan, T.B. : Three Models of Preview Control. IEEE Trans. Human Factors in Electronics HFE-7 (1966), pp. 91 - 102.

[11] Weir, D.H. and McRuer, D.T. : Measurement and Interpretation of Driver/Vehicle System Dynamic Response. Human Factors 15 (1973), pp. 367 - 378.

[12] Wierwille, W.W., Gagné, G.A. and Knight, J.R. : An Experimental Study of Human Operator Models and Closed-Loop Analysis Methods for High-Speed Automobile Driving. IEEE Trans. Human Factors in Electronics HFE-8 (1967), pp. 187 - 201.

MODELLING THE HUMAN OPERATOR OF SLOWLY RESPONDING SYSTEMS USING LINEAR MODELS.

W. Veldhuyzen

Man-Machine Systems Group
Dept. of Mechanical Engineering
University of Technology
Delft.

1. Introduction

From the forties onwards, much attention has been paid to manual control problems. A number of useful models such as the cross-over model [1, 2], and the optimal control model [3] have been developed. The very large part, however, of these studies, is concerned with the control of relatively fastly responding systems, e.g. aircraft and cars, of which the time constants are of the same order as the time constants of the human neuromuscular system.
Since about ten years, there is a tendency to focus more and more the attention on the analysis of human behavior in the supervision of automated systems [4]. More than in the manual control of fastly responding systems, in the supervision of large scale systems are monitoring and failure detection, state estimation, prediction and decision making important aspects of human behavior, on which many investigations will have to be done. Also psychological factors such as attention, motivation, play an important role in the study of supervisory control.
For this reason, describing function techniques, which have proved to be useful in studying problems on manual control of fastly responding systems, are likely not very suitable to the study of problems with slowly responding systems. It should be noted, that a clear distinction between manual and supervisory control cannot be made. Just as in the case of manual control the human supervisor is part of a closed-loop: he receives information about the system, and at discrete times he changes the set-points of the automatic controllers or even switches over from automatic to manual control. In the manual control of a scarlar, slowly responding system it is observed that the human operator behaves in a discrete way [5, 6]: A helmsman of a large ship like a supertanker does not change the rudder position continuously but only at discrete times. The slowly responding character of these ships will be experienced as a kind of supervisory control of a scalar system. Although monitoring, state estimation and decision making are important in this situation and also many psychological factors influence the helmsman's behavior, it will be shown in this paper that the describing function techniques still can be useful in analyzing the control behavior of the helmsman steering a large ship.

2. Experiments

Tests with a maneuvering simulator have been used to analyze the helmsman's control behavior [6]. This simulator is extensively described by Brummer and Van Wijk [7]. In general, the helmsman's task may be considered to be a pursuit tracking task, where the input signal or test signal (the headings ordered) consists of a series of steps of randomly chosen amplitudes and durations. The ordered heading $\psi_d(t)$ has been displayed by means of an alphanumerical display, the actual heading $\psi(t)$ has been presented by means of a compass as normally done.
A time history of the test signal, the subjects had to attempt to follow, is shown in Fig. 1.

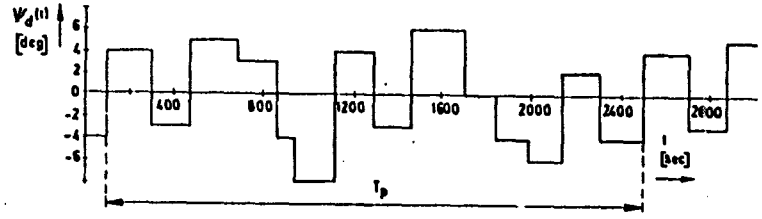


Figure 1 : Time history of the test signal.

The signal is periodic. A test consisted just of one period of forty minutes with a randomly chosen starting point. The choice of this test duration was based on the experience that making the tests too long, the subjects become less motivated at the end of the test, whereas the observation time must be long enough to obtain reliable estimates of the model parameters. Tests have been performed using the signal with the amplitudes as shown in Fig. 1, and with amplitudes twice as large as shown. In the first case the signal is indicated by TS S, in the last case by TS L.

To control the ship's rudder position $\delta(t)$ a steering wheel was provided of which the position is denoted by $\delta_d(t)$.
To simulate the ship's behavior a model describing the dynamics of the ship has to be chosen. A simple model, suitable for this purpose is the following one [8]:

$$T_s \ddot{\psi}(t) + a_1 \dot{\psi}(t) + a_2 [\dot{\psi}(t)]^3 = K_s \delta(t). \qquad (1)$$

The model consists of a nonlinear first order differential equation in the rate of turn, $\dot{\psi}(t)$. The coefficient T_s is related to the ship's moment of inertia with respect to a vertical axis through the center of mass; the coefficient K_s is related to the effective moment which can be exerted on the ship's hull by the rudder; and finally the coefficients a_1 and a_2 are related to the damping. When a_1 is smaller than zero, the ship is directionally unstable, which means that it starts turning to either starboard or port when the rudder is kept amidships. To simulate the steering gear, a first order differential equation has been used, where the rudder angular velocity is limited. In this way the following equations are obtained:

$$T_\delta \dot{\delta}(t) + \delta(t) = \delta_d(t) \; ; \qquad (2a)$$

$$|\dot{\delta}(t)| < \dot{\delta}_m, \qquad (2b)$$

where T_δ is a time constant, and $\dot{\delta}_m$ is the maximum rudder angular velocity.

In this paper, tests with two very large ships, e.g. supertankers, are analyzed. The parameters of these ships with respect to the model (Eqs 1 and 2) are listed in Table 1.

Table 1 : The parameters of the model used to simulate the ships.

Ship	parameters model					
	T_s sec	K_s sec^{-1}	a_1 -	a_2 $(\frac{sec}{deg})^2$	T_δ sec	$\dot{\delta}_m$ $\frac{deg}{sec}$
I	250	-.05	1	5	1	3
II	250	-.05	-1	5	1	3

During these tests no disturbances, such as waves or wind effects, have been introduced in the simulations. That means that the ships were sailing in calm sea.
Four subjects, trainees of the School of Navigation at Amsterdam, were used to analyze the helmsman's behavior. They are indicated by S1, S2, S3 and S4, respectively. None of them was very experienced in steering ships larger than 10,000 tons. To become familiar with the dynamic behavior of large ships, each subject controlled about one hour the unstable ship before starting the experiments.

3. Analysis of the experiments

To identify the helmsman's describing function with respect to each test, several methods are available, which can be divided into two main groups [9]:

- Methods without a-priori knowledge.
- Methods with a certain a-priori knowledge.

In the case that no a-priori knowledge about the system to be identified is available, the identification should be achieved on the basis of general methods such as the determination of Bode or Nyquist plots from the analysis of deterministic test signals or spectral density functions of stochastic processes.

For instance, in a closed loop, the human operator describing function denoted by H(ν) can be determined by the following well-known relation:

$$H(\nu) = \frac{S_{uy}(\nu)}{S_{ue}(\nu)} \; , \qquad (3)$$

where $S_{uy}(\nu)$ and $S_{ue}(\nu)$ are the cross-spectral density functions with respect to the system input u(t), the human operator output y(t) and the error signal e(t), being the difference between system input and controlled element output.
The coherency spectrum, defined as:

$$\Gamma_{uy}(\nu) = \left[\frac{|S_{uy}(\nu)|^2}{S_{uu}(\nu)\, S_{yy}(\nu)}\right]^{1/2} \; , \qquad (4)$$

indicates to what extent the signals u(t) and y(t) are linearly correlated. In Fig. 2 an estimated squared coherency spectrum $\hat{\Gamma}^2_{\psi_d \delta_d}(\nu)$ of a test with the stable ship is shown.

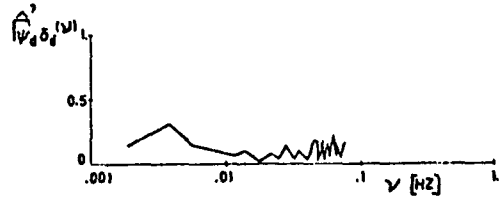


Figure 2 : Estimated squared coherency spectrum $\hat{\Gamma}^2_{\psi_d \delta_d}(\nu)$ of a test with the stable ship.
Subject: S1; Testsignal: TS S.

From this figure, it can be seen that the coherency between the signals $\psi_d(t)$ and $\delta_d(t)$ is rather small, which means that $\delta_d(t)$ is more or less uncorrelated with $\psi_d(t)$ in the frequency range shown. This result corresponds with the fact that for each test the estimated spectral density functions $S_{\psi_d \psi_d}(\nu)$ and $S_{\psi_e \psi_e}(\nu)$, as well as the estimated cross spectrum $|S_{\psi_d \psi_e}(\nu)|$ show only very slight differences in the frequency range observed [6].
Based on these results it can be concluded that the feed back loop does not contain components with frequencies higher than about .01 Hz. It should be noted that the region of interest is only the low frequency range. However, only a few data points can be estimated in this range, since the number and position of data points are determined by the duration of a test, the observation time [10]. This means that the test durations were too short to obtain reliable estimates of the spectra and also of the helmsman's describing function at low frequencies.

When a-priori knowledge is available, for instance the structure of the describing function, the parameters can be determined. Of course, the model obtained in this way should describe the helmsman's control behavior in the frequency range where the spectral analyses of the records did not provide the desired information: the low frequency range. Starting with the simplest human operator model, given by McRuer [1]

$$H_h(j\omega) = K_h \frac{T_1\, j\omega + 1}{T_2\, j\omega + 1} \cdot e^{-j\omega\tau} \qquad (5)$$

taking into account, that slowly responding systems are considered, Eq. (5) can be simplified to Eq. (6):

$$H_h(j\omega) = K_h \frac{T_1\, j\omega + 1}{T_2\, j\omega + 1}, \qquad (6)$$

where the time delay has been neglected because of the slowly responding character of the ship. By assuming that the cross-over model may be applied, it follows that (neglecting again the time-delay)

$$H_h(j\omega).H_s(j\omega) = K_h \frac{(T_1\, j\omega + 1)}{(T_2\, j\omega + 1)} \cdot \frac{K_s}{j\omega(T_s\, j\omega + 1)} = \frac{K}{j\omega} \qquad (7)$$

where the dynamic behavior of the ship has been approximated by Nomoto's first order model [11]. Hence

$$H_h(j\omega) = K_h\,(T_1\, j\omega + 1). \qquad (8)$$

Comparing this model with the linear model used by Stuurman [12]

$$H_h(j\omega) = K_h \frac{(T_1\, j\omega + 1)}{(T_2\, j\omega + 1)}, \qquad (6)$$

it may be expected that the model based on the cross-over model (Eq. 8) has a rather large part of its output power at higher frequencies.
To investigate the influence of the lag term both models have been used to analyze the helmsman's control behavior.

The parameters of the two linear models (Eqs 6 and 8) were estimated as shown in Fig. 3. The upper loop represents the experimental loop with the maneuvering simulator, the lower loop is a simulation of ship and helmsman on a hybrid computer. This method to estimate unbiased model parameters was chosen to be able to analyze also the usefulness of other models, including non-linear models [6].

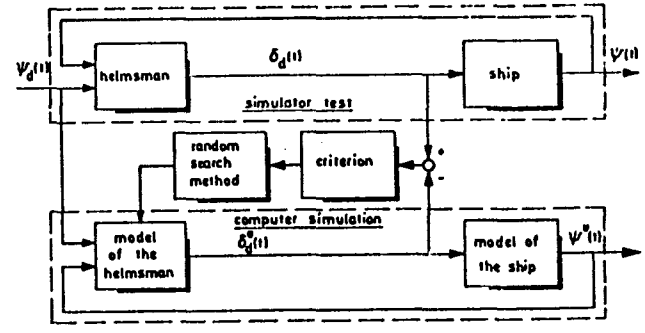


Figure 3 : Estimation of the parameters of the helmsman's models.

The criterion applied to optimize the model parameters was the following one:

$$E_{|\delta|} = \frac{\int_0^T |\,\delta_d(t) - \delta_d^*(t)\,|\,dt}{\int_0^T |\,\delta_d(t)\,|\,dt} \cdot 100\% . \qquad (9)$$

This criterion was preferred to a quadratic criterion

$$E_{\delta}2 = \frac{\int_0^T \left[\delta_d(t) - \delta_d^*(t)\right]^2 dt}{\int_0^T \left[\delta_d(t)\right]^2 dt} \cdot 100\% \qquad (10)$$

as it was expected that the absolute value criterion could be calculated more accurately on the hybrid computer than the quadratic criterion due to the accuracy of the analogue components. However, in literature mostly a quadratic criterion is used, which enables a direct computation of the unknown parameters, if no time-delays are involved [9]. To be able to compare the results to be obtained by minimizing $E_{|\delta|}$ with data given in literature, also $E_{\delta}2$ has been calculated.
Besides the quantity $E_{|\delta|}$ and $E_{\delta}2$, also the following quantities have been computed:

$$E_{\psi}2 = \frac{\int_0^T \left[\psi(t) - \psi^*(t)\right]^2 dt}{\int_0^T \left[\psi(t)\right]^2 dt} \cdot 100\% , \qquad (11)$$

and

$$E_{|\psi|} = \frac{\int_0^T |\,\psi(t) - \psi^*(t)\,|\,dt}{\int_0^T |\,\psi(t)\,|\,dt} \cdot 100\% . \qquad (12).$$

The last two quantities indicate the correspondance between the time histories of the actual heading of the ship $\psi(t)$ steered by the helmsman and those generated by the ship model $\psi^*(t)$ steered by the model of the helmsman.

4. Results

In the Tables 2 and 3 the results of the parameter optimization with the two models are given. These tables provide information about the parameter values determined and the criterion values related to these parameters.
In Fig. 4 some typical time histories are shown of the heading $\psi(t)$ and the steering wheel position $\delta_d(t)$ as well as of the output of the linear model with three parameters (Eq. 6) $\delta_d^*(t)$ and that of the ship model $\psi^*(t)$.

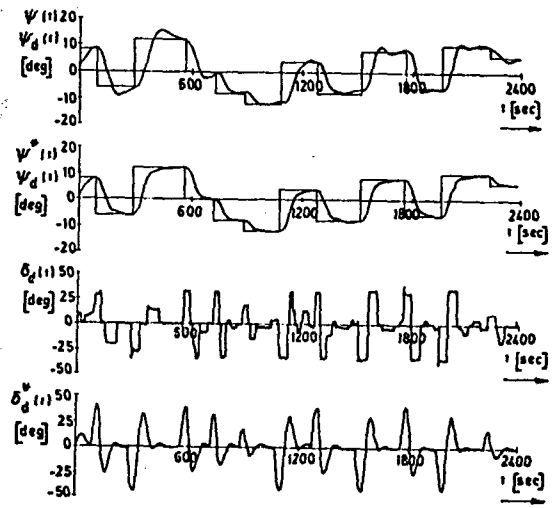


Figure 4 : Typical time histories of the actual signals $\psi_d(t)$, $\psi(t)$ and $\delta_d(t)$ compared with the output of the three parameter model $\delta_d^*(t)$ and the output of the ship model $\psi^*(t)$:
Subject S1; stable ship; TS L.

Table 2 : Results of the parameter optimization with the two parameter model (Eq. 8).

test conditions			parameter values		criterion values			
			K_h	T_1	$E_{\lvert\delta\rvert}$	E_{δ^2}	$E_{\lvert\psi\rvert}$	E_{ψ^2}
Ship	TS	Subj.	-	sec	%	%	%	%
I	S	S1	4.8	32.8	83	60	19	4
		S1	1.2	98.5	88	81	49	26
		S3	3.8	31.6	86	73	-	-
		S4	3.7	41.9	85	68	18	5
	L	S1	5.0	29.5	65	42	-	-
		S1	5.4	28.5	77	58	19	5
		S2	3.8	31.5	73	51	25	8
		S3	3.7	26.5	77	60	33	13
		S3	2.3	43.1	75	57	23	6
II	S	S1	4.5	33.1	69	52	22	6
		S1	5.5	28.1	81	56	32	14
		S2	3.2	45.0	84	72	23	8
		S2	1.6	142.8	95	87	42	21
	L	S1	4.9	33.7	73	56	-	-
		S2	1.0	104.6	85	77	36	15
		S2	1.0	132.7	93	82	-	-

Table 3 : Results of the parameter optimization with the three parameter model (Eq. 6).

test conditions			parameter values			criterion values			
			K_h	T_1	T_2	$E_{\lvert\delta\rvert}$	E_{δ^2}	$E_{\lvert\psi\rvert}$	E_{ψ^2}
Ship	TS	Subj.	-	sec	sec	%	%	%	%
I	S	S1	5.5	46.2	8.1	74	46	19	4
		S1	2.8	89.4	21.6	75	61	32	11
		S3	6.5	73.2	19.6	76	61	-	-
		S4	4.5	64.3	9.1	78	59	20	6
	L	S1	4.3	48.3	11.3	49	22	-	-
		S1	8.2	46.7	13.4	55	30	16	4
		S2	5.1	49.6	11.8	56	30	13	2
		S3	4.1	46.9	11.2	67	46	14	3
		S3	3.1	69.4	19.1	64	43	14	3
II	S	S1	5.1	43.2	6.2	60	36	18	4
		S1	5.9	37.6	7.8	71	43	30	13
		S2	3.2	67.8	18.8	62	42	21	6
		S2	2.6	136.9	23.8	81	67	37	15
	L	S1	4.6	53.9	11.9	59	35	-	-
		S2	1.9	94.9	21.5	74	57	15	3
		S2	2.1	106.8	31.2	75	50	-	-

To interprete these results in terms of the cross-over model, it is necessary to linearize the nonlinear ship model. To this end the mean value and the variance of the rate of turn for each of the tests are computed. The results of these computations are shown in Table 4.

Table 4 : Computed mean value and variance of the ship's rate of turn for each test.

test conditions			mean value	variance
Ship	TS	Subj.	deg/sec	$(deg/sec)^2$
I	S	S1	.0034	.0046
		S1	-.0025	.0024
		S3	-.0025	.0030
		S4	-.0002	.0040
	L	S1	-.0041	.0116
		S1	.0029	.0092
		S2	.0015	.0130
		S3	.0014	.0141
		S3	-.0031	.0122
II	S	S1	.0045	.0060
		S1	-.0036	.0052
		S2	-.0007	.0046
		S2	-.0045	.0025
	L	S1	-.0035	.0138
		S2	-.0040	.0130
		S2	-.0020	.0095

From this table it was concluded that the rate of turn was such small during all the experiments that the influence of the nonlinear term of the ship model, $a_2[\dot{\psi}(t)]^3$, can be neglected without making too large errors.
Also the dynamics of the steering gear are nonlinear. However, by linearizing this behavior a first order system is obtained, of which the time constant can be estimated to be three through four seconds. This time constant is small in relation to the ship time constant T_2. Hence, the influence of the steering gear on the stability of the system helmsman-ship will be small too. A more exact estimation of the time constant has not been executed for this reason.
Based on the parameter values given in the Tables 2 and 3, and the linearized ship model the cross-over frequencies and the phase margins have been determined for each test by means of Bode plots. The results are shown in Table 5.

Table 5 : The cross-over frequencies and the phase margins with respect to the two helmsman's models.

test conditions			two param. model		three param. model	
Ship	TS	Subj.	ω_c rad/sec	ϕ deg	ω_c rad/sec	ϕ deg
I	S	S1	.03	53	.05	50
		S1	.03	79	.05	35
		S3	.03	51	.08	31
		S4	.02	51	.05	54
	L	S1	.03	42	.04	45
		S1	.03	49	.05	38
		S2	.03	50	.05	42
		S3	.03	46	.04	44
		S3	.02	52	.04	39
II	S	S1	.03	37	.04	42
		S1	.03	32	.04	33
		S2	.03	44	.04	27
		S2	.05	77	.05	24
	L	S1	.03	37	.05	34
		S2	.02	52	.03	30
		S2	.02	58	.04	20

5. *Discussion and conclusions*

The structure of the model has been based on data given in literature, in particular McRuer's cross-over model. According to this model the helmsman adapts his control behavior to the controlled element dynamics in such a way that the necessary conditions required for a good closed loop response are fulfilled, that means that the closed loop system has to be stable and well-damped with a high amplitude ratio of the open loop frequency response $|H_p H_e|$ for frequencies of the input bandwidth and a low amplitude ratio outside this range.
From Table 5 it can be concluded that for each of the two models these requirements are fulfilled: all the parameters found correspond to a stable and mostly well-damped system, even in the case of the unstable ship.
In the literature on automatic steering of ships, a large number of studies can be found on the design of autopilots. Koyama [13] used a controller with a transfer function equal to three parameter helmsman's model (Eq. 6) to control an unstable ship, with a time constant of about 269 sec. He has found that for T_1 = 100 sec, T_2 = 12.5 sec, and K_p between 1 and 4 sec, a stable closed loop response is obtained, of which the performance is very acceptable in a wide range of ship speeds. In accordance with the cross-over model the parameter values found (Table 3) agree very well with the values given by Koyama.

The criterion values with respect to the steering wheel position, $E_{|\delta|}$ and $E_{\delta}2$, indicate how well in terms of these criteria the models describe the helmsman's control behavior. From the Tables 2 and 3, it can be concluded that the two parameter model yields a much poorer description than the three parameter model. The criterion values $E_{\delta}2$ with respect to the latter range from 30 until 60%, with an average of about 45%, with only a few exceptions. This means that this model provides a rather poor average description of the helmsman's control behavior. In particular in the case of the three parameter model, the correspondance between the heading of the ship as steered by the model closely approximates the heading of the ship as steered by the helmsman. The criterion values $E_{\psi}2$ are less than 10%, with a few exceptions. Although the output of the model of the helmsman's behavior sometimes differs from the actual output, the heading of the ship generated by the model is invariably a good fit due to the very low pass filtering properties of the ships.

Summarizing the following conclusions can be drawn:

- To identify the helmsman's describing function at very low frequencies, the test duration must be very long; however, after about 45 minutes the subjects are getting tired, and unwanted effects will occur. Hence, the identification of the helmsman's describing function by means of e.g. spectral analysis is not possible in the situation studied in this paper.
- The results obtained show that the cross-over model can be useful in the analysis of human behavior in controlling slowly responding systems.
- The description of the helmsman's control behavior by means of a model consisting of a gain and a lead term is poor; by adding a lag term a better description can be obtained.
- The heading of the ship steered by the two parameter model matches the heading of the ship steered by the helmsman rather well. In the case of the three parameter model a good description has been achieved.

6. Acknowledgement

The author gratefully acknowledge the contribution of Mr. J.F. Zegwaard of the Hybrid Computer Centre, who executed the model optimizations on the hybrid computer, and the valuable assistance in data processing of Mr. R.E. Schermerhorn.

7. References

1. McRuer, D.T.; Jex, H.R.,
A review of quasi-linear pilot models.
IEEE-trans. on Human Factors in Electronics,
Vol. HFE-8 (1967), No. 3 (Sept.), pp. 231-249.

2. McRuer, D.T.; Krendel, E.S.,
Mathematical models of human pilot behaviour.
Report: NATO-AGARD, No. 188, 72 p.

3. Kleinman, D.L.; Baron, S.; Levison, W.H.,
A control theoretic approach to manned-vehicle systems analysis.
IEEE-trans. on Autom.Control
Vol. AC-16 (1971), No. 6 (Dec.), pp. 824-832.

4. Proceedings of an International Symposium on Monitoring Behavior and Supervisory Control, Berchtesgaden F.R. of Germany, Ed. by T.B. Sheridan and G. Johannsen.
NATO Conference Series, Series III: Human Factors,
Plenum Press New York: 527 p.

5. Cooke, J.E.,
Human decisions in the control of a slow response System.
Diss.: Oxford, 1965, 403 p.

6. Veldhuyzen, W.,
Ship manoeuvring under human control: analysis of the helmsman's control behaviour.
Ph.D. thesis, Department of Mechanical Engineering,
Delft University of Technology, 1976, 104 p.

7. Brummer, G.M.A.; Wijk, W.R. van,
The ship manoeuvring and research simulator of the Institute TNO for Mechanical Constructions, Delft.
Report: Delft, Inst. TNO for Mech. Constr., 1970, No. 8133/1, 32 p.

8. Norrbin, N.H.,
On the design and analysis of the zig-zag test on base of quasi-linear frequency response.
Proc. Tenth Int. Towing Tank Conf. 1963, pp. 355-374.

9. Lunteren, A. van; Stassen, H.G.,
Annual Report 1969 of the Man-Machine Systems Group.
Report: Delft, Dept. of Mech. Engineering, 1970, No. WTHD 21, 102 p.

10. Jenkins, G.M.; Watts, D.G.,
Spectral Analysis and its Applications.
Holden Day, San Francisco, 1969.

11. Nomoto, K.; Taguchi, T.; Honda, K.; Hirano, S.,
On the steering qualities of ships.
I.S.P. Vol. 4 (1957), No. 35 pp. 354-370.

12. Stuurman, A.M.,
Modelling the helmsman: A Study to Define a Mathematical Model Describing the Behaviour of a Helmsman Steering a Ship along a straight course.
Report: Delft, Inst. TNO for Mech. Constr., 1969, No. 4701, 59 p.

13. Koyama, T.,
Some notes on the auto-pilot of an unstable ship.
Report: Delft, Shipbuilding Laboratory, 1971, No. 327, 23 p.

EFFECTS OF SIMULATED SURFACE-EFFECT SHIP MOTIONS ON CREW HABITABILITY

Overview of Program and Selected Visual/Motor Results for Phase I, IA*

Henry R. Jex, Richard J. DiMarco, Warren F. Clement,
Jeffry R. Hogge, and Steven H. Schwartz
Systems Technology, Inc., Hawthorne, California

Abstract

As part of the U.S. Navy's development program for large high-speed surface-effect-ships (SES), a series of four related motion simulations were performed during 1973 to 1976, starting with the Six-Degree-of-Freedom Motion Base at NASA's Marshall Space Flight Center, then on the ONR/HFR Ship Motion Generator (MoGen) at Goleta, California.

The first portion of this paper gives an overview of the scope of these intense programs, which involved close cooperation of several agencies including: USN Bureau of Ships, USN Bureau of Medicine, Naval Ship Research Development Center, Human Factors Research, Inc., and Systems Technology, Inc., Oceanics Inc., and NASA. Comprehensive measurements were made on a variety of cognitive, visual/motor, and habitability tasks related to the SES missions. Both objective and subjective data were taken, as well as medical observations, EEG's while sleeping, and stress-hormone excretion rates.

The motions simulated were those predicted via a complex math model for an early version of a typical 2000 Ton SES of roughly 300 ft × 100 ft (100 m × 30 m) dimensions traveling at speeds of 40-80 kts (21-41 m/sec) in Sea States from below 3 to 5. RMS heave accelerations of the wideband waveforms ranged from .08 to .28 g at characteristic frequencies from .5 to 1.2 Hz, with varying content in the motion sickness frequency range near .3 Hz. Sample motions are shown for the primary sea state/ship's speed conditions with statistical parameters given for the others.

The second portion of the paper describes some of the visual motor and habitability rating data, primarily from Phases I and IA, which were run during 1974 on the ONR/HFR MoGen. The tasks covered here are the: Critical-Instability Task, ("ECM Tracking") Lock Opening Task, Keyboard (calculator) Operation, and Head-Point-of-Regard motions fixating dials while seated. In addition subjective ratings of motion sickness (kinetosis) and of motion effects and interference are summarized, and are correlated with motion intensity.

*This work was sponsored by the U.S. Navy Surface Effect Ship Project (SESP, Project PMS-304), at Carderock, MD.; under the technical direction of Warren Malone and Lt. Cdr. M. J. Vickery (Royal Navy).

DRIVER STEERING DYNAMICS MEASURED IN A CAR SIMULATOR UNDER A RANGE OF VISIBILITY AND ROADMARKING CONDITIONS

R. Wade Allen and Duane T. McRuer
Systems Technology, Inc.
Hawthorne, California

Abstract

A simulation experiment was conducted to determine the effect of reduced visibility on driver lateral (steering) control. The simulator included a real car cab and a single lane road image projected on a screen six feet in front of the driver. Simulated equations of motion controlled apparent car lane position in response to driver steering actions, wind gusts, and road curvature. Six drivers experienced a range of visibility conditions at various speeds with assorted roadmarking configurations (mark and gap lengths).

Driver describing functions were measured and detailed parametric model fits were determined. A pursuit model employing a road curvature feedforward was very effective in explaining driver behavior in following randomly curving roads. Sampled-data concepts were also effective in explaining the combined effects of reduced visibility and intermittent road markings on the driver's dynamic time delay. The results indicate the relative importance of various perceptual variables as the visual input to the driver's steering control process is changed.

INTRODUCTION

Automobile steering control is a dynamic task that is performed by the driver in order to establish and/or maintain the vehicle on a specified pathway in the presence of inputs such as crosswinds and roadway curvature. The motions of an automobile in response to steering actions and aerodynamic disturbances can be described in terms of differential equations, transfer functions, etc. (Ref. 1), and it is logical to attempt to describe the driver in similar terms.

In the research described herein the motivation for a dynamic description and measurement of the driver was twofold: 1) to determine and quantify the effect of adverse visibility on driver perception of the cues required for steering control; and 2) to determine those changes in driver behavior that contribute to degraded performance under conditions of reduced visibility. An understanding of these effects may then suggest countermeasures to vehicle control problems associated with adverse visibility.

PERCEPTION AND DRIVER/VEHICLE SYSTEM GUIDANCE REQUIREMENTS

Given that road delineation features (e.g., dashed lines) are detected, the driver's perceptual processes can then extract information for vehicle control. Adverse visibility conditions can reduce visual range (Ref. 2), and the question to be considered here is how reduced visibility might affect driver perception of vehicle path and subsequent control actions.

An abstraction of the driver's perceptual task is illustrated in Fig. 1, a perspective view of a single lane bounded by dashed lines. With forward motion in a straight line the driver's visual scene appears to expand from a perspective vanishing point at infinity. Theories have been advanced for a focus or center of expansion perception of motion (Ref. 3). On straight roads the center of expansion is the only point in the visual field that is apparently stationary, and it would provide a direct cue for the car's path angle. Thus, when forward view is reduced by adverse visibility, direct perception of path angle is denied the driver according to the focus of expansion theory. This theory has some problems, however. As Gordon (Ref. 4) notes, for curved paths the center of expansion lies at the center of curvature, which is at right angles to the path of the vehicle. Furthermore, Palmer (Ref. 5)

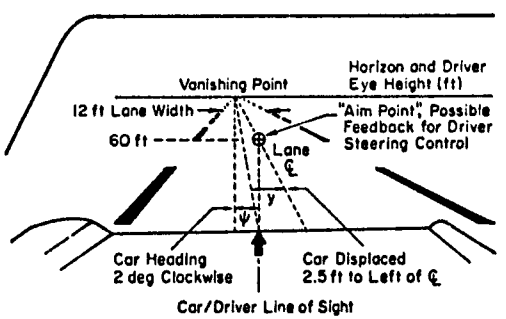


Figure 1. Driver's Perspective View of a Single Delineated Path Illustrating an Aim Point Control Law

has found that the center of expansion in visual fields expanding at various constant rates of expansion can only be perceived within 1-6 degrees of visual angle, which is much too coarse for vehicular control.

In a perspective motion field the streamers themselves play a more important role in the views of Calvert (Ref. 6), who emphasized their role in both directional and longitudinal control of aircraft on the final approach, and Gordon (Ref. 4), who considered terrestrial vehicles. The streamer theory states in essence that the driver perceives motion from objects in the visual field streaming across his field of view. Although the streamers emanate from the center of expansion, Gordon believes that it is the streamers themselves, particularly those provided by roadway boundaries and lane markings, that underlie the directional cue rather than the center of expansion. He notes that all parts of the visual field, road borders, and lane markers move when the wheel is turned but no one part is essential for tracking, and that the driver responds to a total situation (a Gestalt concept), not to isolated or ranked cues. Streamer perception should be fairly robust in the face of reduced visibility, although adverse visibility (i.e., rain, fog) would eliminate many subtle cues (e.g., road roughness, edge texture), particularly those available outside foveal vision where contrast sensitivity and acuity degrade (Ref. 7).

Control theory analysis and research into land vehicle steering control have identified cues that must be perceived either explicitly or implicitly in order to give good, stable performance. The car's position relative to the delineated path is the most obvious of these. Various studies have also demonstrated that heading or path angle is essential to achieving stable control (as reviewed in Ref. 8). Thus, properly weighted components proportional to lateral position and heading must be present in the driver's steering wheel deflection if the car's path is to be regulated in the lane.

One intuitively appealing model for driver lateral control involves steering inputs based on an aim point down the road as illustrated in Fig. 1. The aim point angle is one way to combine lateral position and preview-range-weighted heading into a single control quantity. The dynamics of this simple control model, among others, have been analyzed previously (Ref. 9). For an

aim point at a distance, x_a, a look-ahead or preview time constant dependent on vehicle speed, U_o, can be defined as follows:

$$T_a = \frac{x_a}{U_o} \qquad (1)$$

McLean (Ref. 10) has reviewed a number of driving experiments involving variations in restricted forward view and vehicle speed which found preview times (T_a) of 2 sec or greater. The results were quite variable, however, and it would be difficult to decide on an average or typical preview time constant.

If there is a preferred look-ahead distance or time constant, then restricted visual range due to adverse visibility could interfere with this cue, and visual ranges shorter than the preferred look-ahead distance would be expected to deteriorate performance. Lane position and heading cues do not necessarily have to be perceived at a combined aim point, however. Referring again to Fig. 1, simple geometric analysis shows that for small angles the car's heading angle deviations with respect to the lane appear as horizontal translations of the visual scene. For car lateral position deviations with respect to the lane the road appears to rotate about its vanishing point at the horizon. Thus heading and lateral positions are separately available from the perspective view if a sufficient segment of this view is visible.

DRIVER/VEHICLE SYSTEM DYNAMIC MODEL

To gain further insight into driver perceptual requirements, consider the detailed driver/vehicle system dynamic model illustrated in Fig. 2. Here the vehicle model gives heading angle and lateral lane deviations (ψ and y) in response to driver steering commands (δ_w). The driver develops steering commands based on his perception of lane position and heading angle errors, plus an additional term proportional to perceived road curvature. The y and ψ perceptions are basically involved in regulation-only driver control, which is handled in a compensatory fashion. The added curvature term is a pursuit or feedforward element needed to account for driver behavior on curved roads. It basically assumes the driver inserts an open-loop steering wheel command

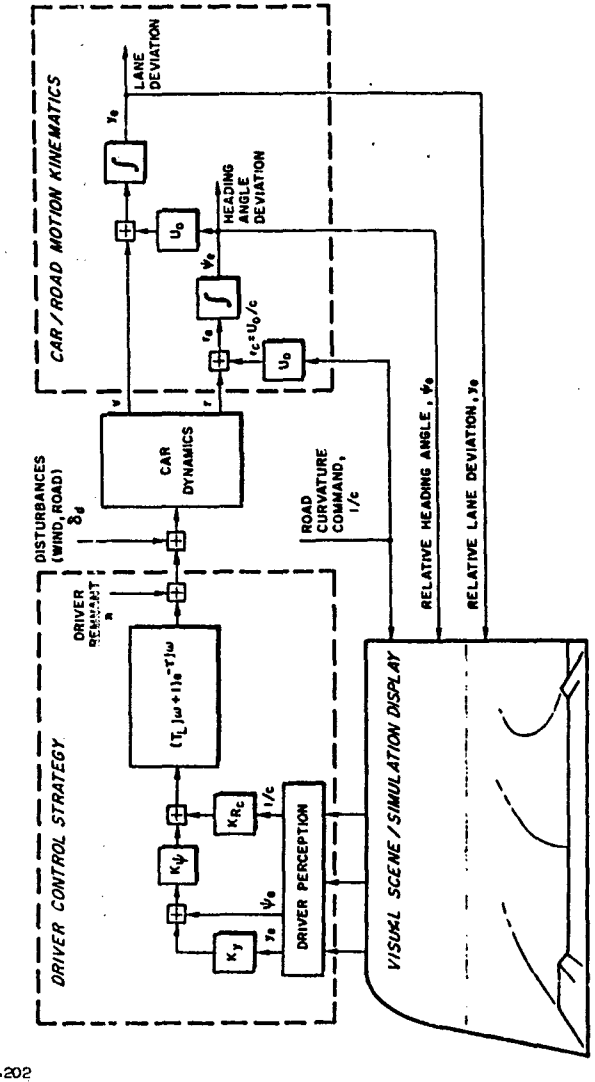


Figure 2. A Driver/Vehicle Dynamic System Model for Analyzing Adverse Visibility Effects on Steering Control

proportional to perceived path curvature. Some anticipation or driver lead (T_L) is applied to these perceptions to offset vehicle lag, and a time delay penalty (τ) is incurred by the driver due to basic neuromuscular characteristics and perceptual processing load. A final component of the driver's steering action is composed of remnant which is basically noise or random variation in the driver's output uncorrelated with perceptual inputs.

The regulation or error correcting portion of the Fig. 2 model, involving only lane position and heading error feedbacks (y_e and ψ_e, respectively) has been shown to have good, stable control properties (Ref. 11) and to be consistent with experimental measurements (Refs. 9 and 12). This mode of control is termed compensatory in that the driver is "compensating" for disturbance caused errors. In the case of a wind disturbance the driver has no perception or preview of the disturbance, and must wait for the disturbance to affect the vehicle's motion before responding.

When following a curving road (a path command) with sufficient visual range, the driver has the opportunity to preview or anticipate the desired path. With a visual segment large enough to permit adequate perception of the road's curvature, the driver can achieve a pursuit mode of control behavior and very nearly duplicate the commanded path. This is simply accomplished by the driver because in steady state the curved path followed by a car is nearly directly proportional to front wheel angle (Ref. 13), and the vehicle lags are well-learned and can be anticipated. Thus, the driver merely steers with actions directly proportional to perceived road curvature, sufficiently advanced in time to offset vehicle lag. Disruption of the curvature cue will degrade pursuit performance, however, which is a possibility with various combinations of adverse visibility and delineation as discussed previously.

COMPENSATORY VERSUS PURSUIT BEHAVIOR

Consider now the control and performance implications of compensatory and pursuit behavior. Given the model structure of Fig. 2 and nominal driver parameters obtained in this study under good visibility at 30 mph, we have analyzed the dynamic implications of curvature perception. Referring to Fig. 2, consider the driver/vehicle system response due to a command path

input. The commanded path causes an equivalent heading rate input, r_c, to be applied to the system. Driver steering action should then create vehicle yaw rates, r, which are nearly equal to the commanded heading rate so as to give a small heading rate error, r_e, which is the difference between the command input and the vehicle's motion.

To illustrate the potential improvement in performance between pursuit and compensatory driving we can consider the describing function relating heading rate error, r_e, to heading rate input command, r_c. Figure 3 shows Bode plots of this driver/vehicle system ratio as the curvature perception parameter (driver's "pursuit gain"), K_{R_C}, is increased. The compensatory baseline curve ($K_{R_C} = 0$) is based on a representative set of driver/vehicle data, and the other curves simply indicate the effect of $|r_e/r_c|$ when the additional driver control pathway represented by K_{R_C} is added. At low frequencies the describing function amplitude shows that errors are less than the original input, while at very high frequencies they may be somewhat greater. Note in particular that for the ideal no-lag vehicle an optimum

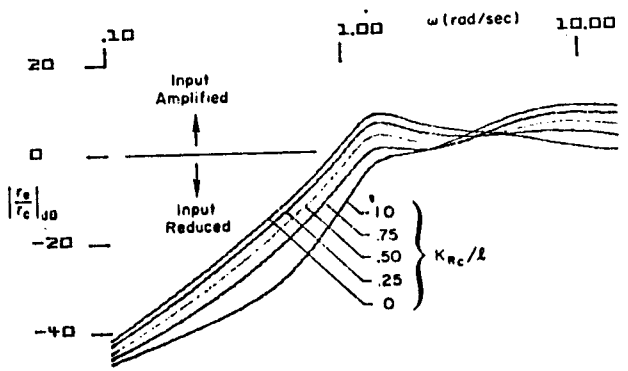


Figure 3. Effects of Variations in the Curvature Perception Gain on the Heading Error Rate to Heading Rate Input Transfer Function

value* of $K_{R_C}/\ell \doteq 1$, errors in the frequency region of 0.3-1.0 rad/sec give a reduction in error of about 15 dB or a factor of greater than 5 times! At any given frequency, lane dispersions are directly proportional to heading rate errors and thus the curvature perception in the above frequency region would reduce lane dispersions by more than a factor of 5.

ADVERSE VISIBILITY EFFECTS

The Fig. 2 model can serve as the basis for some observations about driver visual perception requirements and potential effects of degraded visibility. Consider first the driver's use of the aim point concept illustrated in Fig. 1. Here reductions in visual range under adverse visibility conditions can eliminate the cues required to directly perceive the aim point. In this case the driver can extrapolate from the available cues or, alternatively, separately perceive lateral and heading error deviations. In either case, however, the driver is faced with an increased perceptual load. Past research has shown that increased perceptual load leads to increases in time delay (τ) and noise or remnant (Refs. 14 and 15). These effects should increase with decreased visual range.

When reduced visual range interferes with direct perception of the aim point, the lane delineation configuration then should become an important factor. Consider Fig. 1 with restricted preview. The driver needs adequate information to perceive ψ_e, y_e, and road curvature. If several delineation elements are visible, or single elements are of sufficient length, these variables should be directly perceivable. If element length is reduced, however, so that path direction is not readily indicated by a single element, then two components are needed to define direction and three to indicate curvature. In terms of the Fig. 2 model, a visual segment which contains at least three elements is needed for development of the K_{R_C} feedforward, while at least two elements are needed for ψ_e to be estimated. Thus the driver/vehicle system dynamics will depend strongly on the dimensions of the visual segment. As

*For the essentially neutral steer car of this study the steady-state turn radius is equal to the wheelbase divided by the front wheel angle (Ref. 13), so the curvature perception gain should be equal to the car's wheelbase, in this case 9.25 feet.

it is reduced, performance on curves will be degraded first (K_{R_C} reduced), followed by deterioration in lane position control. For segmented delineation the driver's input information also becomes perceptually intermittent as the visual segment is reduced, and intermittency has been shown to lead to increased time delay and remnant in the human operator (Ref. 19).

Besides providing insight into degraded visibility effects, the Fig. 3 model also serves as a paradigm for data measurement and analysis. The driver control strategy parameters in Fig. 3 can be determined through Fourier analysis techniques (Ref. 9) during driving tests involving regulation against disturbances, and following winding roads. This technique was used here as described further on to measure the perceptual/behavioral effects of adverse visibility, simultaneously during a single realistic task (in situ as it were), rather than requiring a series of artificial tests to isolate each effect.

EXPERIMENTAL METHODS

Simulation Setup

A fixed-base driving simulator was used to test the concepts discussed previously. The physical arrangement of the simulator is illustrated in Fig. 4. Simulator details have been described previously (Ref. 16). The simulator has a high quality, wide angle video projection display of roadway markings as illustrated. Display perspective and motion were correctly represented with respect to the driver's eye position, and the electronic display generator was designed to allow a variety of delineation configurations and visibility conditions. Apparent road motion relative to the cab was controlled by driver steering, acceleration, and braking actions through equations of motion mechanized on an analog computer.

Delineation configuration and visual range could easily be controlled from an experimenter's console. The mark and cycle lengths of dashed delineation lines could be independently selected in discrete steps. The range extent of the visual segment could be set continuously from zero to a maximum display generator range of 300 ft. Visual range was controlled by an electronic intensity function which smoothly decreased display delineation line luminance as

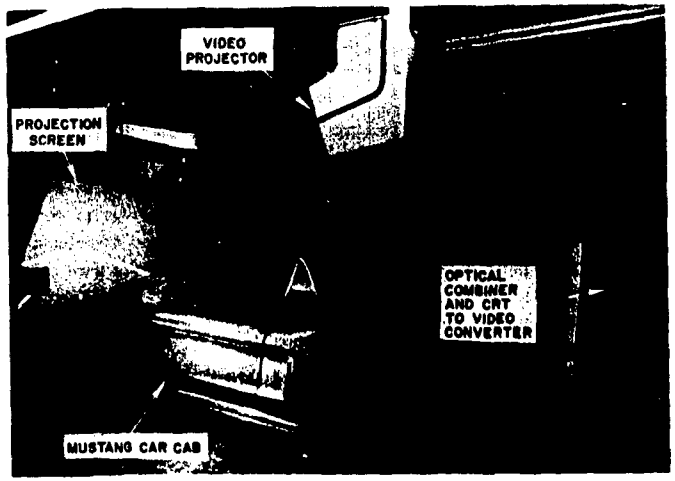


Figure 4. Driving Simulator Physical Arrangement

a function of distance down the road. Desired visual range was set subjectively as described later. This gave the desired physical results directly and minimized the need to control or account for all the subtle photometric and subjective effects which determine threshold contrast.

Experimental Design

Based on an exploratory series of tests the test matrix shown in Fig. 5. was evolved. The matrix includes the important range of the three major variables of interest: visibility range, delineation configuration, and speed. The visibility range extends from close to the minimum possible for steering control (35 ft) out to a distance beyond that required for good control. Speed variations are covered from very slow to the current nationwide speed limit. Delineation configuration applies to a single lane delineated with left and right boundaries and covers a California standard with

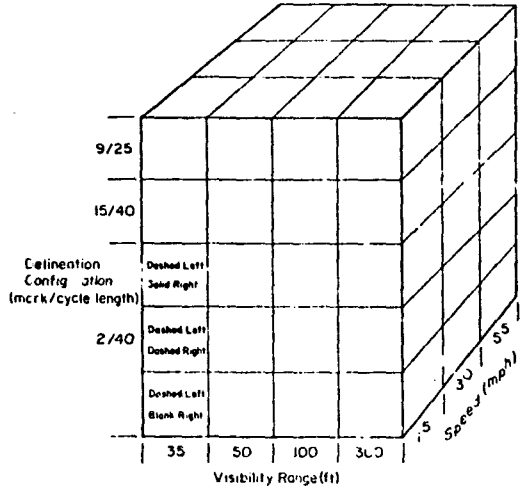


Figure 5. Simulator Test Condition Matrix

9 ft marks and 25 ft cycle (Ref. 17), the national standard recommendations of 15 ft marks and 40 ft cycles (Ref. 18), and a very short element spaced at 40 ft meant to simulate retroreflectors which individually offer no directional cues. A further variation was applied to the right boundary of the retroreflector delineation which included either a solid, dot (reflector) or blank (no right edge line) configuration. A solid edge line would presumably improve performance over that with dotted elements, while the lack of any edge line at all would degrade performance under adverse visibility. It was impractical to run all combinations of the factors shown in Fig. 5, so the combinations listed in Table 1 were selected to span the major dimensions, with emphasis on combinations likely to show degraded performance (i.e., higher speeds, shorter visibility ranges, and shorter delineation elements).

Typical examples of visibility and configuration conditions are illustrated in Figs. 6 and 7. In Fig. 6 it is apparent how reduced visual range affects lateral position, heading, and curvature cues. In Fig. 7 the effects

TABLE 1. SIMULATION EXPERIMENTAL CONDITIONS

VISIBILITY RANGE (ft)	SPEED (mph)	CONFIGURATION		CONFIGURATION VISIBILITY PARAMETER, C_v (Dimensionless)	PLOTTING SYMBOL[b]
		MARK/CYCLE LENGTH (ft)	LEFT/RIGHT LANE EDGE CONFIGURATION[a]		
300	30	15/40	D/D	0.09	▷
100	30	15/40	D/D	0.28	▼
100	55	15/40	D/D	0.28	▼
100	30	2/40	D/D	0.54	■
100	55	2/40	D/D	0.54	■
50	15	9/25	D/D	0.44	○
50	30	9/25	D/D	0.44	●
50	55	9/25	D/D	0.44	●
50	15	15/40	D/D	0.55	◁
50	30	15/40	D/D	0.55	◀
50	55	15/40	D/D	0.55	◀
50	15	2/40	D/D	1.08	◇
50	15	2/40	D/S	1.08	◇
50	15	2/40	D/N	1.08	◇
50	30	2/40	D/D	1.08	◆
50	30	2/40	D/N	1.08	◆
50	30	2/40	D/S	1.08	◆
35	15	15/40	D/D	0.79	△
35	30	15/40	D/D	0.79	▲

[a] D - dashed; S - solid; N - none.

Categorized as follows:
- Solid symbols used for 15/40 (MUTCD Standard)
- Dashed symbols used for 9/25 (California Standard)
- Dotted symbols used for 2/40 (MUTCD Standard cycle length with short marks or bits)

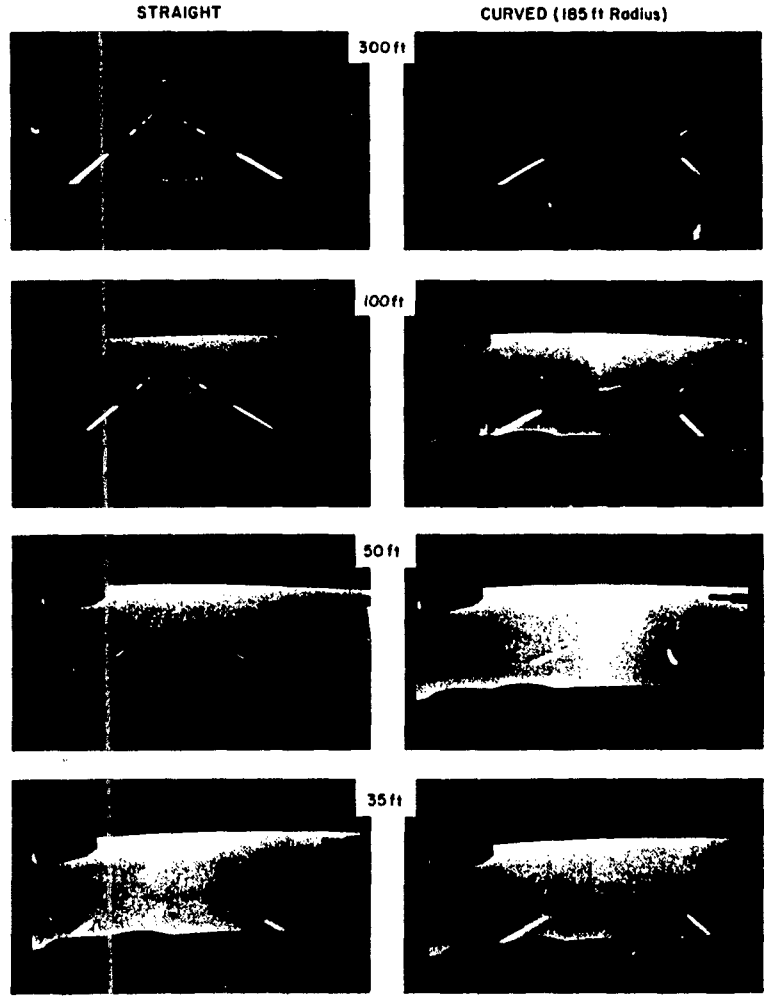


Figure 6. Federal Standard Delineation Striping Under Various Levels of Reduced Visibility

P-202

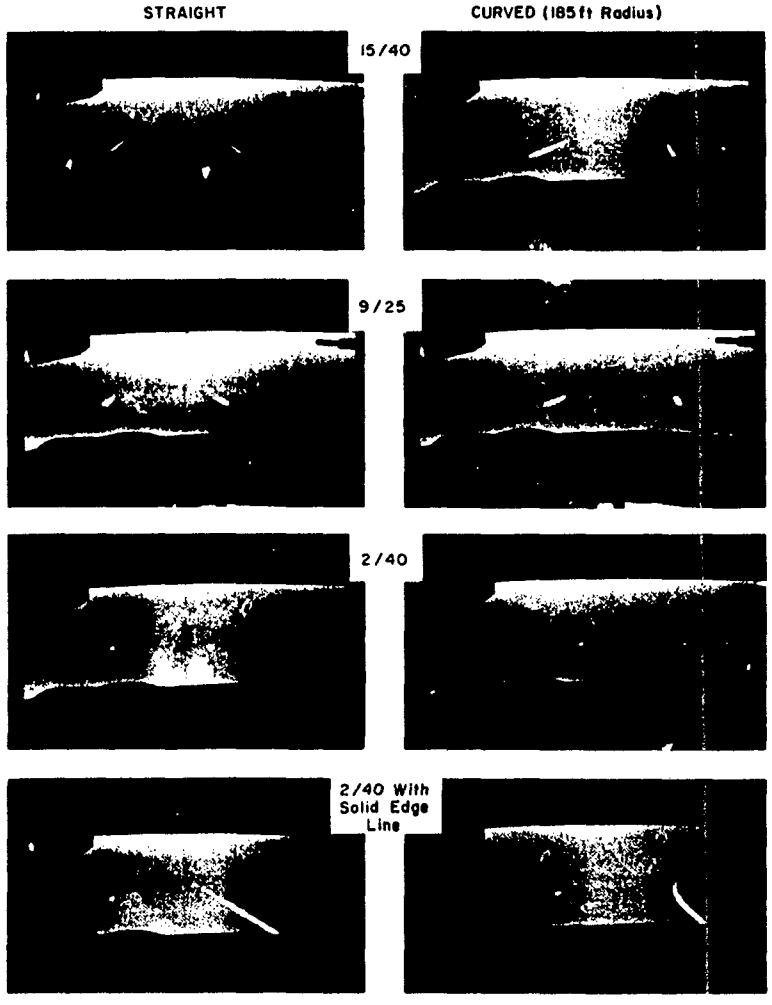


Figure 7. Delineation Configuration Variations Under 50 ft Visual Range

P-202

of the various delineation configurations under reduced visual range are apparent; the decreased directional cues with shorter delineation elements, and the vast improvement with a solid edge line.

Based on the exploratory experiment results a configuration visibility parameter listed in Table 1 was developed to quantify the combined perceptual effects of delineation configuration and visual range. As illustrated in Fig. 8, the configuration visibility parameter has two components. The first, $(x_g + x_o)/x_v$, is related to the number of delineation elements visible. For $(x_g + x_o)/x_v > 1$ we see from Fig. 8 that delineation elements close to the car are obscured by the hood before the next element becomes visible down the road (hood visibility obstruction in the simulator was 19 ft). The second

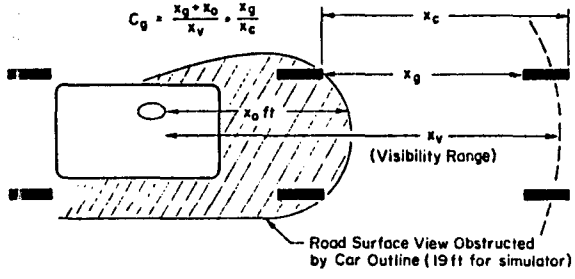


Figure 8. Configuration Visibility Parameter for Quantifying the Combined Effects of Delineation Configuration and Visual Range

factor, x_g/x_c, roughly quantifies the proportion of available information when delineation elements are visible. Thus, the configuration parameter gives large values for poor visibility, large gaps, and large proportions of gap-to-cycle length ratio, and driver performance would be expected to deteriorate under these conditions. The configuration parameter quantifies therefore both the visual segment and intermittency aspects of the driver's visual scene. This parameter thus far only accounts for symmetrical delineation, however, and we will have to consider the performance effects to determine what influence the right edge line variations will have.

Procedures

Six licensed drivers with normal vision were selected as test subjects. Background on the subjects is given in Table 2. Prior to the formal testing/data-gathering experiments, each subject was given a brief introductory session that consisted of driving the test scenario twice through each baseline configuration. The purpose of these sessions was twofold: 1) to transition subjects' skills of everyday automobile driving to the fixed-base simulator environment; and 2) to introduce the subject to the general nature of the test plan and procedures. These initial sessions were intentionally short in order to minimize training efforts and avoid overlearning.

TABLE 2. SUBJECT BACKGROUND

SUBJECT	SEX	AGE	DRIVING EXPERIENCE (IN YEARS)	EDUCATION
A	M	29	17	B.S.
B	M	21	5	A.A.
C	M	33	17	A.S.
D	M	29	13	A.A.
E	F	41	24	H.S.
F	F	17	1	H.S.

Typically, each subject drove the entire test scenario in two or three days of concentrated testing. Each day was broken into four test sessions of one to one and a half hours each separated by a rest period. At the beginning and end of each day, the 300 ft baseline configuration was tested. At the beginning and end of each session, a baseline configuration was administered, either the 300 ft or 50 ft visibility condition. Baseline conditions were also periodically interspersed within the sessions. All configurations in the experiment were given to the subjects in a pseudo-random order. At no time was a subject aware of what conditions would be driven next.

In order to control for within- and between-subject variations in contrast thresholds and equipment variations, visibility distance was individually set

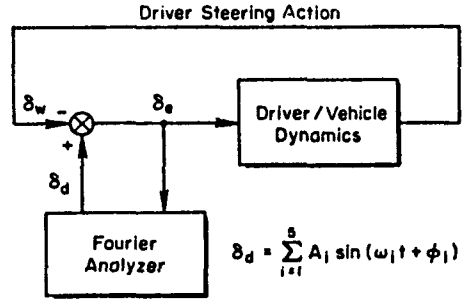
for each condition according to the following procedure. The experimenter would initially set the visibility range, then ask the subject to position a line, which appeared across the roadway, to the point at which the delineation disappeared. The line position was controlled with a ten-turn potentiometer and was returned to zero between estimates. The experimenter would repeat this procedure several times, readjusting the visibility range between estimates in an iterative procedure until the desired visibility range, as indicated by the subject, was achieved.

Tasks and Measures

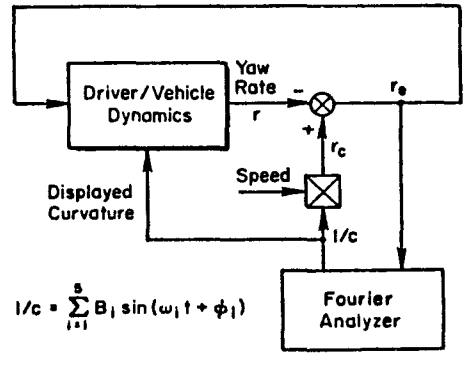
Describing function and performance measurements were obtained under two task conditions for each experimental condition. One task required regulating against a random wind gust-like disturbance added in at the steering signal input to the vehicle equations of motion as illustrated in Fig. 2. This task required compensatory control behavior as the roadway was straight and the disturbance could not be observed other than in its effect on vehicle motions. A second task involved following a winding road which allowed for pursuit control behavior if the visual scene provided for adequate curvature perception. Curvature commands for this task were added into the equations as shown in Fig. 2 in addition to curving the displayed roadway.

The method for obtaining driver describing function data is shown in Fig. 8. A Fourier analyzer (Ref. 19) generated a sum of sine waves input (Table 3) that was injected into the system as either a command or a disturbance, and received back another system quantity which was subsequently Fourier analyzed at each of the input frequencies ω_i.

As noted in Fig. 8 the actual quantities used to compute the equivalent driver/vehicle open-loop describing function depend on the task input. For the winding road command input case, where pursuit behavior is possible, the error (r_e) to input (r_c) describing function is computed and then transformed to give an equivalent open-loop transfer function r/r_e. For the compensatory, wind gust disturbance input the equivalent open-loop transfer function is found from operations on the δ_e/δ_d ratio, as described in Ref. 9.



a) Wind Gust Disturbance Input



b) Winding Road Command Input

Figure 8. Driver Describing Function Measurement Technique

TABLE 3. INPUT AMPLITUDES AND FREQUENCIES

FREQUENCY (rad/sec)	WIND GUST AMPLITUDES (EQUIVALENT FRONT WHEEL ANGLE, deg)	ROAD CURVATURE AMPLITUDES (INVERSE RADIUS OF CURVATURE, ft^{-1})
0.188	0.172	2.07×10^{-3}
0.503	0.172	2.06×10^{-3}
1.25	0.172	1.98×10^{-3}
3.	0.172	1.66×10^{-3}
6.28	0.086	0.56×10^{-3}
rms	0.251	2.79×10^{-3}

After the describing function data are developed, an optimal identification routine is used to find driver parameters for the Fig. 2 system model that will give a good match to the measured describing function data.

As an example consider the data illustrated in Fig. 9 for the two baseline visibility conditions. The measured describing functions were averaged across six subjects, and the describing function fits match the data rather well. The characteristic effect of the curve perception parameter K_{RC} is apparent in both cases in comparing the compensatory (wind gust disturbance) and pursuit (winding road command) tasks.

RESULTS

Model parameters for both steering tasks over a number of visibility and delineation configuration conditions are compared in Table 4. The most apparent consistent effect in the complete model parameters seems to be the reduction in curvature perception (K_{RC}) with increased configuration visibility parameter C_V. This relationship is plotted in Fig. 10. Although the effects are not as neat as we might hope, the tendencies are quite clear: curve perception gain decreases with increasing configuration visibility parameter, and decreasing speed. Changes due to C_V are undoubtedly associated with the amount of curvature information provided by the visual segment on a

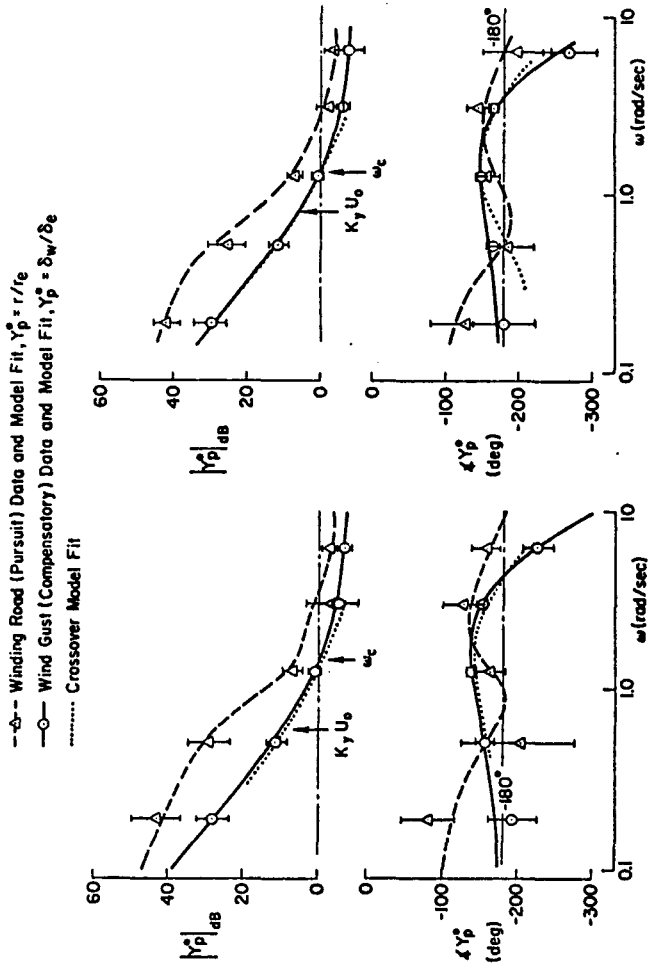


Figure 9. Baseline Condition Driver/Vehicle Describing Functions and Transfer Function Model Fits

TABLE 4. DYNAMIC RESPONSE PARAMETERS

EXPERIMENTAL CONDITION				PURSUIT (WINDING ROAD COMMAND)						COMPENSATORY (WIND GUST DISTURBANCE)							
				MODEL FIT PARAMETERS						COMPLETE MODEL FIT PARAMETERS						CROSSOVER CHARACTERISTICS	
VISIBILITY RANGE (ft)	MARK/CYCLE LENGTH (ft)	CONFIGURATION VISIBILITY PARAMETER, C_v	SPEED, U_o (mph)	$\frac{K_{RC}}{I}$	K_Y (ft^{-1})	K_ψ	T_L (sec)	τ (sec)		K_Y (ft^{-1})	K_V	T_L (sec)	τ (sec)	$\frac{1}{K_Y U_o}$ (sec)	ω_c (rad/sec)	Φ_m	
200	15/40	0.09	30	0.857	0.0357	0.174	0.247	0.422		0.022	0.195	0.471	0.526	1.03	1.46	39.02	
100	2/40	0.54	30	0.903	0.0286	0.136	0.164	0.450		0.0355	0.105	0.652	0.600	0.64	1.15	28.23	
50	9/25	0.44	15	0.795	0.159	0.220	0.276	0.490		0.0130	0.274	0.342	0.691	3.5	0.855	45.02	
50	9/25	0.44	30	0.959	0.0114	0.260	0.152	0.551		0.0319	0.141	0.553	0.628	0.71	1.38	26.44	
50	9/25	0.44	55	1.005	0.0098	0.224	0.154	0.577		0.0258	0.132	0.477	0.471	0.48	2.21	13.76	
50	15/40	0.55	15	0.772	0.0436	0.156	0.310	0.518		0.0113	0.315	0.109	0.541	4.0	0.785	42.86	
50	15/40	0.55	30	0.856	0.0218	0.244	0.241	0.540		0.0251	0.169	0.485	0.620	0.91	1.30	27.58	
50	15/40	0.55	55	0.979	0.0033	0.0769	0.769	0.604		0.0231	0.133	0.530	0.552	0.54	2.20	13.33	
50	2/40	1.08	30	0.571	0.0819	0.0854	0.453	0.58		0.0301	0.154	0.389	0.652	0.76	1.12	16.29	
35	15/40	0.79	15	0.141	0.0605	0.355	0.447	0.587		0.0116	0.225	0	0.328	3.92	0.697	28.38	
35	15/40	0.79	30	0.825	0.0143	0.26	0.115	0.529		0.0176	0.206	0.352	0.672	1.29	1.36	24.64	

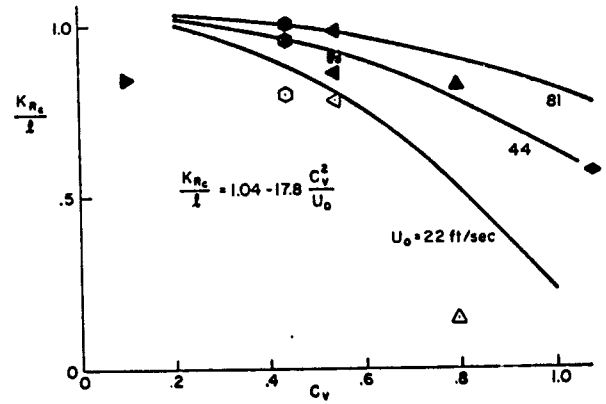


Figure 10. Effects of Adverse Visibility and Delineation Configuration on the Curvature Perception Parameter

static basis as discussed earlier. The speed effect is most likely related to the dynamic or "streaming" effect of delineation elements moving in the visual field, with higher speeds inducing a stronger perception of curvature.

Aside from the curve perception parameter, K_{RC}, the large number of parameters which comprise the complete model data in Table 4 exhibit covariations that make it difficult to focus the effects of adverse visibility on any one particular parameter. To permit more generalization we need a simpler summary of the driver's dynamic response that would nonetheless have properties pertinent to closed-loop driver/vehicle performance. For this purpose an "extended" crossover model is appropriate. For the equivalent compensatory open-loop transfer function of Fig. 11 this has the form:

$$\frac{\delta}{\delta_e} = \frac{s + (1/U_o K_Y)}{s^2} \omega_c e^{-\tau_e s} \qquad (2)$$

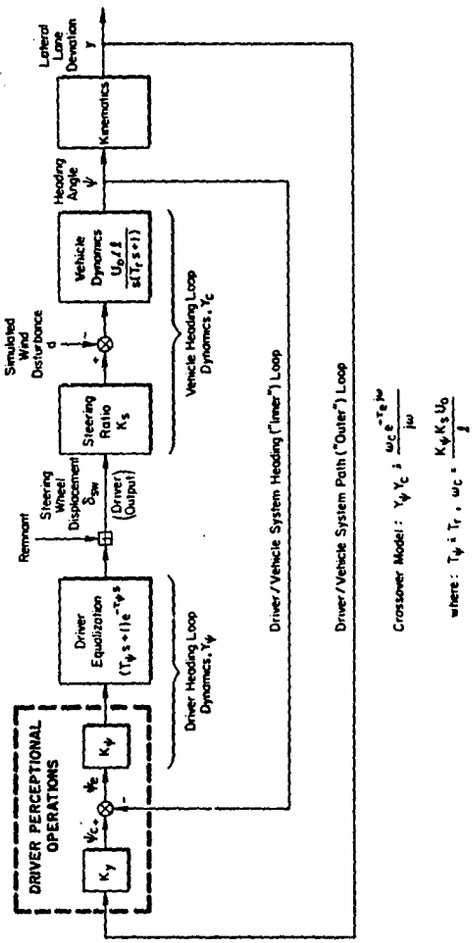


Figure 11. Simplified Crossover Model Interpretation of Driver/Vehicle Steering Control

This relationship is based on the equivalent driver/vehicle transfer function developed in Ref. 9. The equivalent time delay τ_e combines the high-frequency phase properti of the driver (lead, neuromuscular lags, transport delay) and vehicle (basically the heading response dynamics). The crossover frequency, ω_c, combines the driver and vehicle heading gains which can be expressed by the useful approximation (Ref. 2)

$$\omega_c = \frac{K_\psi K_s U_o}{\ell} \tag{3}$$

where K_ψ is the driver's heading gain from the complete model, K_s is the steering ratio (which has arbitrarily been set to unity for all the K_ψ parameters reported herein), and ℓ is the vehicle wheelbase. The free s in the denominator of Eq. 2 approximates the wheel input to car heading angle dynamics. The phase lags of the additional high-frequency lag properties are accounted for in τ_e, and typical crossover frequencies are low enough that the high frequency amplitude properties are not important. Finally, the numerator zero at $(U_o K_y)^{-1}$ accounts for the driver's operations on lane position. $U_o K_y$ can be interpreted as an equivalent look-ahead time constant (T_z in Eq. 1) and K_y^{-1} is the corresponding look-ahead distance.

A simple approximation can be used to compute the parameters of Eq. 2. From moderately low to high frequencies, the phase angle of the numerator zero at $(U_o K_y)^{-1}$ can be approximated by an exponential (Ref. 20)

$$\angle \frac{[s + (U_o K_y)^{-1}]}{s} \doteq e^{-\alpha/s} \tag{4}$$

where $\alpha \doteq (U_o K_y)^{-1}$. Using this approximation, the phase of Eq. 2 then can be written as

$$\angle \frac{\delta}{\delta_e} = -\tau_e \omega - \frac{\alpha}{\omega} - \frac{\pi}{2} \tag{5}$$

This equation can now be evaluated at the gain and phase crossover frequencies of the driver/vehicle describing function in order to solve for τ_e and α. At gain crossover frequency the phase angle is equal to π less the phase margin

φ_m, and the phase crossover is defined as the frequency at which the phase angle is equal to π. Therefore, τ_e and α can be computed from the equation:

$$\begin{bmatrix} \omega_c & 1/\omega_c \\ \omega_u & 1/\omega_u \end{bmatrix} \begin{bmatrix} \tau_e \\ \alpha \end{bmatrix} = \begin{bmatrix} (\pi/2) - \varphi_m \\ \pi/2 \end{bmatrix} \qquad (6)$$

Using this relationship a few key dynamic response parameters can be calculated for the driver/vehicle system that are appropriate to the basic closed-loop properties and performance.

Cross-model parameters were fitted to all the wind gust disturbance (compensatory) conditions according to the above procedures, and these are listed in Table 5. The fits are representative of the data as illustrated in Fig. 9. Crossover modal parameters are plotted in Fig. 12. Heading gain, K_ψ, was computed from ω_c data using Eq. 3. There is a small effect of both configuration visibility and speed on K_ψ.

Inverse outer-loop (lane position) gain can be interpreted as an equivalent dynamic look-ahead distance and was computed using the relationship

$$K_y^{-1} \doteq U_o \alpha \qquad (7)$$

The Fig. 11b plot of these data shows a strong dependence on C_v, with shorter distances apparently associated with restricted visual range (i.e., higher C_v values). There is little apparent speed dependence, however, suggesting that look-ahead distance is a more pertinent perceptual variable than the look-ahead time constant of Eq. 1.

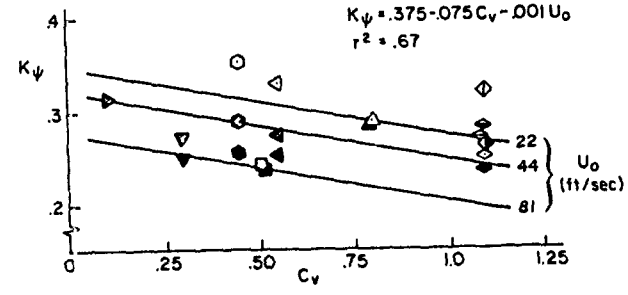
Effective time delay, τ_e, is affected by both speed and configuration visibility parameter C_v. The increase in effective time delay with increase in C_v suggests that perceptual load associated with equalization and extrapolation increases with the reduction in the visual segment. Past research has postulated that τ_e increases are associated with an increase in driver lead equalization requirements and/or with the presentation of sampled information. In this case driver lead is used to offset the vehicle lag. But,

TABLE 5. EXTENDED CROSSOVER MODEL PARAMETERS

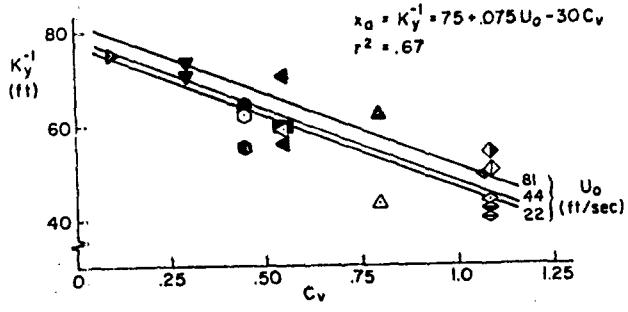
EXPERIMENTAL CONDITION				ω_c (rad/sec)	ω_u	φ_m	τ_e (sec)	$\alpha \doteq U_o K_y$ (rad/sec)	$\frac{1}{U_o K_y}$ (sec)	K_y^{-1} (ft)	K_ψ
VISIBILITY RANGE (ft)	MARK/CYCLE LENGTH (ft)	CONFIGURATION VISIBILITY PARAMETER, C_v	SPEED, U_o (mph)								
300	15/40	0.09	30	1.462	4.290	39.02	0.334	0.587	1.70	75	0.307
100	15/40	0.28	30	1.295	3.758	35.56	0.375	0.601	1.66	73.2	0.272
100	15/40	0.28	55	2.167	4.385	22.48	0.298	1.153	0.867	70	0.249
100	2/40	0.54	30	1.147	3.672	28.25	0.372	0.747	1.34	59	0.241
100	2/40	0.54	55	2.060	3.800	13.76	0.317	1.398	0.715	58	0.237
50	9/25	0.44	15	0.855	3.367	45.02	0.436	0.352	2.84	62	0.351
50	9/25	0.44	30	1.377	3.528	26.44	0.381	0.806	1.24	55	0.289
50	9/25	0.44	55	2.21	3.565	13.96	0.341	1.269	0.788	64	0.254
50	15/40	0.55	15	0.785	3.312	42.86	0.440	0.375	2.67	59	0.330
50	15/40	0.55	30	1.302	3.607	27.58	0.375	0.782	1.28	56	0.274
50	15/40	0.55	55	2.205	3.295	13.33	0.271	1.146	0.87	70	0.253
50	2/40	1.08	15	0.595	2.687	24.78	0.516	0.495	2.02	44	0.250
50*	2/40	1.08	15	0.762	3.240	37.71	0.443	0.438	2.28	50	0.320
50*	2/40	1.08	15	0.643	3.395	26.2	0.416	0.544	1.84	40	0.270
50	2/40	1.08	30	1.117	2.895	16.29	0.436	0.893	1.12	49	0.235
50*	2/40	1.08	30	1.347	3.210	15.44	0.388	1.05	0.952	42	0.283
50*	2/40	1.08	30	1.253	3.705	26.70	0.365	0.812	1.23	54	0.263
35	15/40	0.79	15	0.687	2.918	28.30	0.478	0.516	1.93	43	0.289
35	15/40	0.79	30	1.355	2.897	24.64	0.458	0.704	1.42	62	0.285

*Not used in regression analysis because of non-symmetrical delineation (see Table 1).

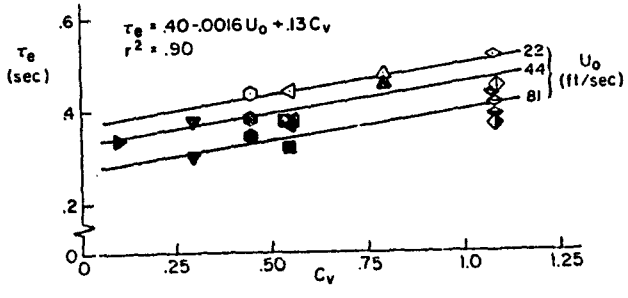
P-202



a) Heading Gain



b) Equivalent Dynamic Look-Ahead Distance



c) Equivalent System Latency

Figure 12. Driver/Vehicle System Extended Crossover Model Parameters for Compensatory Task

the vehicle lag increases with speed whereas the effective time delay decreases with speed, so this "explanation" is in the wrong direction. The sampling interval of the dashed lines is given by

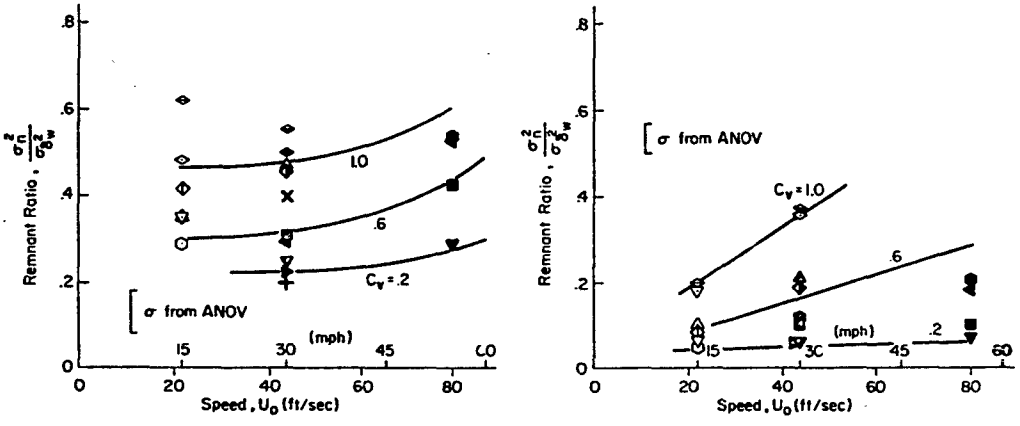
$$T_s = \frac{x_c}{U_o} \tag{8}$$

where x_c is the delineation cycle length. Change in τ_e is proportional to sampling interval (Ref. 15), so the variation of τ_e with both speed and C_v is in the right direction. Accordingly, we attribute the τ_e changes primarily to sampling processes associated with the delineation dashed lines and speed.

The sampling process which affects the driver's time delay should also have some influence on the noise or stochastic component of his steering actions. In Fig. 13 we show the proportion of noise or remnant that is uncorrelated with the driver's actions in counteracting wind gusts or steering along a winding road. There is a tendency for driver noise to increase both with configuration parameter and speed.

The intermittency of delineation apparently affects driver remnant; however, this effect is increased at higher speeds (i.e., higher delineation sample rates) in contrast to the time delay penalty (Fig. 12) which decreased with increasing speed. These two contrasting effects on driver behavior explain the relatively consistent effect of speed on performance under adverse visibility shown in Fig. 14. At low speeds, the slow intermittency of delineation causes appreciable increases in driver time delay which degrades performance; while at high speeds, driver noise increases, which again degrades performance. Also note that for the curve-following data driver noise increases appreciably under the same conditions that led to reduced curvature perception (Fig. 10). Furthermore, the curve-following data show a proportionately greater configuration visibility parameter sensitivity, consistent with the Fig. 14 performance data. A final observation is that the solid edge line reduces driver noise over the dashed or no right lane line cases, which (as with previously discussed data) is consistent with a lower equivalent C_v.

P-202

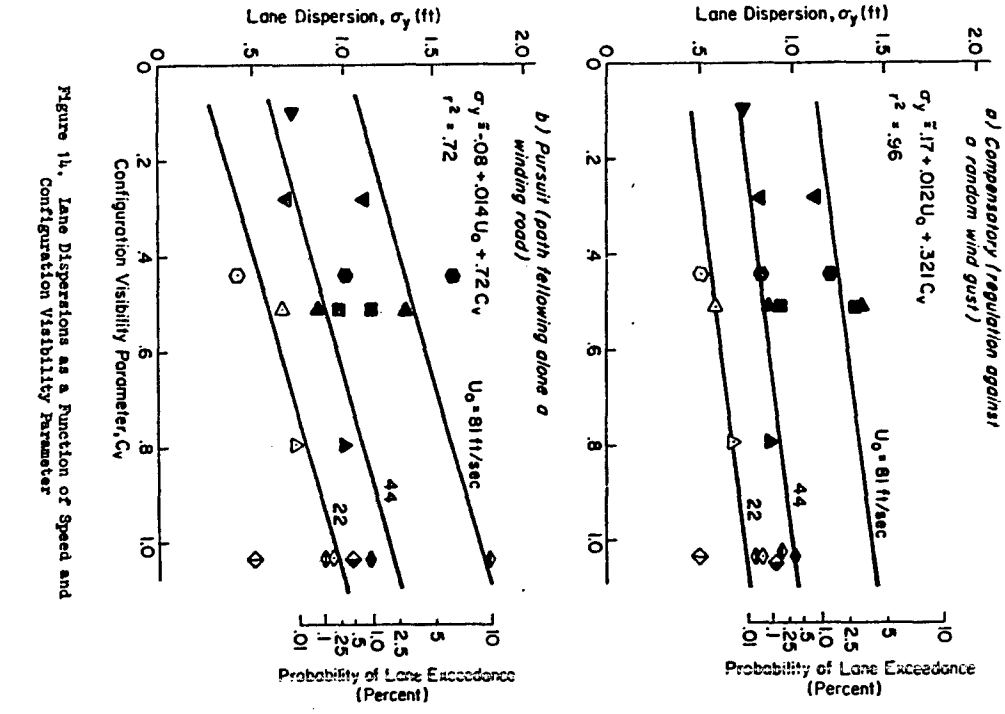


a) Compensatory Control (regulation against a random wind gust)

b) Pursuit Control (path following along a winding road)

Figure 13. Driver Remnant as a Function of Speed and Configuration Visibility Parameter

P-202



a) Compensatory (regulation against a random wind gust)

b) Pursuit (path following along a winding road)

Figure 14. Lane Dispersions as a Function of Speed and Configuration Visibility Parameter

CONCLUDING REMARKS

Adverse visibility restricts the driver's perception of automobile path and motion information required for steering control. These perceptual restrictions can be quantified in terms of the driver's dynamic steering behavior in response to random disturbances and path commands. Combinations of reduced visibility and delineation configuration (i.e., intermittent dashed or dotted lines) tend to induce increased transport delay in the driver and impair his perception of road curvature. Reduced visibility also induces a reduction in equivalent dynamic look-ahead distance (the inverse of lateral position error gain) but does not appear to influence the weighting or gain the driver applies to heading errors.

The above effects appear to be related to the apparent intermittent or sampled nature of delineation under reduced visibility conditions. Driver time delay, τ_e, increases at slower speeds, due to decreased sampling frequency, even though vehicle dynamic lags decrease with speed. This effect induced a somewhat compelling urge in some subjects to speed up in order to increase their information rate, which is a rather insidious phenomenon if true for real-world driving, since it might encourage drivers to maintain speeds with associated stopping distances exceeding their visual range.

Changes in curve perception gain also appear to be related to information sampling, with K_{R_C} decreasing with decreased speed and/or reduction in the amount of perceptual information (i.e., increased C_V). In this case curvature perception is enhanced with speed, which may also be related to "streamer" theories of driver perception where the curved path motion of the car is indicated by the curved motion of visual field elements.

This research has given some insight into dynamics of driver perception, and the role played by road markings used to delineate the commanded pathway. The techniques developed here should be useful for further research in this area and allow various questions to be answered about the required or optimal configurations of roadway delineation.

REFERENCES

1. Segel, L., "Theoretical Prediction and Experimental Substantiation of the Response of the Automobile to Steering Control," in Research in Automobile Stability and Control and in Tyre Performance, London, Institute of Mechanical Engineers, 1957.

2. Allen, R. W., and D. T. McRuer, "The Effect of Adverse Visibility on Driver Steering Performance in an Automobile Simulator," presented at the 1977 SAE International Automobile Engineering Congress and Exposition, Detroit, 28 Feb.-4 Mar. 1977.

3. Gibson, J. J., "What Gives Rise to the Perception of Motion," Psych. Review, Vol. 75, 1968, pp. 335-346.

4. Gordon, D. A., "Perceptual Basis of Vehicular Guidance," Public Roads, Vol. 34, No. 3, Aug. 1966, pp. 53-68.

5. Palmer, E. A., "Experimental Determination of Human Ability to Perceive Aircraft Aimpoint from Expanding Gradient Cues," Aerospace Medical Meeting, San Francisco, Preprint of Scientific Program, May 1969, pp. 176-177.

6. Calvert, E. S., "Visual Judgments in Motion," J. Inst. Navigation, Vol. 7, 1957.

7. Haines, R. F., "A Review of Peripheral Vision Capabilities for Display Layout Designers," Proc. S.I.D., Vol. 16/4, Fourth Quarter, 1975, pp. 238-249.

8. McRuer, D. T., R. W. Allen, D. H. Weir, and R. H. Klein, "New Results in Driver Steering Control Models," Human Factors, Special Issue, forthcoming.

9. McRuer, D. T., D. H. Weir, H. R. Jex, R. E. Magdaleno, and R. W. Allen, "Measurement of Driver/Vehicle Multiloop Response Properties with a Single Disturbance Input," IEEE Trans., Vol. SMC-5, No. 5, Sept. 1975, pp. 490-497.

10. McLean, J. R., and E. R. Hoffmann, "The Effects of Restricted Preview on Driver Steering Control and Performance," Human Factors, Vol. 15, No. 4, Aug. 1973, pp. 421-430.

11. Weir, D. H., and D. T. McRuer, "Dynamics of Driver/Vehicle Steering Control," Automatica, Vol. 6, No. 1, Jan. 1970, pp. 87-98.

12. McRuer, D. T., R. H. Klein, et al., Automobile Controllability — Driver/Vehicle Response for Steering Control. Vol. I: Summary Report. Vol. II: Supporting Experimental Results, DOT HS-801 407 and HS-801 406, Feb. 1975.

13. McRuer, D. T., "Simplified Automobile Steering Dynamics for Driver Control," presented to the SAE Aerospace Control and Guidance Systems Comm. Mtg. No. 35, Palo Alto, Calif., 19-21 Mar. 1975.

14. McRuer, D. T., and E. S. Krendel, Mathematical Models of Human Pilot Behavior, AGARD-AG-188, Jan. 1974.

15. Allen, R. W., W. F. Clement, and H. R. Jex, Research on Display Scanning, Sampling, and Reconstruction Using Separate Main and Secondary Tracking Tasks, NASA CR-1569, July 1970.

16. Allen, R. W., J. R. Hogge, and S. H. Schwartz, "A Simulator for Research in Driver, Vehicle and Environment Interaction," presented at the 56th Meeting of the TRB, Jan. 1977.

17. Traffic Manual, State of California, Business and Transportation Agency, Dept. of Public Works, 1971.

18. Manual on Uniform Traffic Control Devices for Streets & Highways, Federal Highway Admin., 1971.

19. Allen, R. W., and H. R. Jex, "A Simple Fourier Analysis Technique for Measuring the Dynamic Response of Manual Control Systems," IEEE Trans., Vol. SMC-2, No. 5, No. 1972, pp. 638-643.

20. McRuer, D., D. Graham, E. Krendel, and W. Reisener, Jr., Human Pilot Dynamics in Compensatory Systems — Theory, Models, and Experiments with Controlled Element and Forcing Function Variables, AFFDL-TR-75-15, July 1965.

ACKNOWLEDGMENT

This work was supported by the Federal Highway Administration under Contract DOT-FH-11-8824. Dr. Donald A. Gordon of the Traffic Systems Division served as the Contract Technical Manager.

Session IV

MONITORING BEHAVIOR AND SUPERVISORY CONTROL

Chairman: W. B. Rouse

N79-17494

ABSTRACT

Supervisory Dynamic Decision-Making in Multi-Task Monitoring and Control

by

M. K. Tulga and T. B. Sheridan
Massachusetts Institute of Technology

As computers become smaller, cheaper and more sophisticated, the tasks of the human pilot and ground controller are changing radically from that of being a continuous controller in one or a few control loops to that of being a monitor of many separate tasks, or a supervisory-coordinator of semi-automated subsystems. Human operators of nuclear and other process control plants are undergoing a similar change in role. Quantitative models by which to describe and predict behavior are lacking, however, and therefore need to be developed. This paper describes preliminary research in one such modeling effort.

In the experimental paradigm, a number of "tasks" are represented simultaneously on a computer display as rectangles of varying heights (representing relative "value density" of given tasks) and widths (representing task "durations"). Tasks appear at random times and places and move at fixed velocity toward a "deadline" at the right-hand margin. The subject's objective is to "attend" to one task at a time (hold a cursor, by means of a graphics tablet, successively to different rectangles) and thus cause that task's width to collapse at a uniform rate, hopefully to disappear before the deadline is reached and the opportunity time therefore lost. The reward earned is the aggregate of reduction in areas of all tasks. "Tasks" may be clustered in groups with additional delays imposed for switching the cursor from one group to another. The subject obviously tries to concentrate on the most rewarding tasks, but may lose time by changing the cursor to unimportant tasks or by shifting to a different task group. The experiment is implemented on an Imlac graphics terminal coupled to an Interdata display, and provides a wide variety of parameter combinations.

An optimal decision control model has been developed, which is based primarily on a dynamic programming algorithm which looks at all the available task possibilities, charts an optimal trajectory, and commits itself to do the first step (i.e., follow the optimal trajectory during the next time period), and then iterates the calculation. A Bayesian estimator has also been included which estimates the tasks which might occur in the immediate future and provides this information to the dynamic programming routine.

Preliminary trials comparing the human subject's performance to that of the optimal model show a great similarity, but indicate that the human skips certain movements which require quick change in strategy. We are planning to rerun the model under a variety of parameter combinations (principally, rate of discounting future data and/or preview span, and response time delay) to solve the inverse optimal problem. That is, for what parameters do the human data and the optimal model match?

Further, we will attempt to show how the experimental paradigm (and model) correspond to future flight management tasks.

MODELING HUMAN DECISION MAKING BEHAVIOR IN SUPERVISORY CONTROL +

M. K. Tulga
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Mass. 02139 U.S.A.

T. B. Sheridan
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Mass. 02139 U.S.A.

As computers become smaller, cheaper and more sophisticated, the tasks of the human pilot and ground controller are changing radically from that of being a continuous controller in one or a few control loops to that of being a monitor of many separate tasks, or a supervisory-coordinator of semi-automated subsystems. Human operators of nuclear and other process control plants are undergoing a similar change in role. Quantitative models by which to describe and predict behavior are lacking, however, and therefore need to be developed. This paper describes preliminary research in one such modeling effort.

In the experimental paradigm, a number of "tasks" are represented simultaneously on a computer display as rectangles of varying heights (representing relative "value density" of given tasks) and widths (representing task "durations"). Tasks appear at random times and places and move at fixed velocity toward a "deadline" at the right-hand margin. The subject's objective is to "attend" to one task at a time (hold a cursor, by means of a graphics tablet, successively to different rectangles) and thus cause that task's width to collapse at a uniform rate, hopefully to disappear before the deadline is reached and the opportunity time therefore lost. The reward earned is the aggregate of reduction in areas of all tasks. "Tasks" may be clustered in groups with additional delays imposed for switching the cursor from one group to another. The subject obviously tries to concentrate on the most rewarding tasks, but may lose time by changing the cursor to unimportant tasks or by shifting to a different task group. The

+ The authors gratefully ackrowledge support from the National Aeronautics and Space Administration, Grant No. NSG-2118.

experiment is implemented on an Imlac graphics terminal coupled to an Interdata display, and provides a wide variety of parameter combinations.

An optimal decision control model has been developed, which is based primarily on a dynamic programming algorithm which looks at all the available task possibilities, charts an optimal trajectory, and commits itself to do the first step (i.e., follow the optimal trajectory during the next time period), and then iterates the calculation. A Bayesian estimator has also been included which estimates the tasks which might occur in the immediate future and provides this information to the dynamic programming routine.

Preliminary trials comparing the human subject's performance to that of the optimal model show a great similarity, but indicate that the human skips certain movements which require quick change in strategy. We are running the model under a variety of parameter combinations (principally, rate of discounting future gain and/or preview time, and response time delay) to solve the inverse optimal problem. That is, for what parameters do the human data and the optimal model match?

INTRODUCTION

With the advent of computer technology which resulted in increased computing power and speed, and in its ever increasing availability due to decreasing costs and size (Ref. 1), the role of the human operator in manned systems is changing from that of a continuous "in-line" controller to that of a monitor/supervisor, (Ref. 2). He is therefore asked to supervise multiple dynamic processes, each of which is controlled continuously by a hierarchy of "intelligent" machines. Higher-level automation is already a commonplace practice in space- and aircraft technology (Ref. 3), and is now gaining acceptance in industrial plants as well.

PARADIGM

We can model the situation of supervisory control as the decision-maker (DM) monitoring and controlling a system of processes as shown in Figure 1.

Note that in Figure 1 the number of processes that the DM can supervise is limited mainly because of the limitations in information transmission to the DM through the displays.

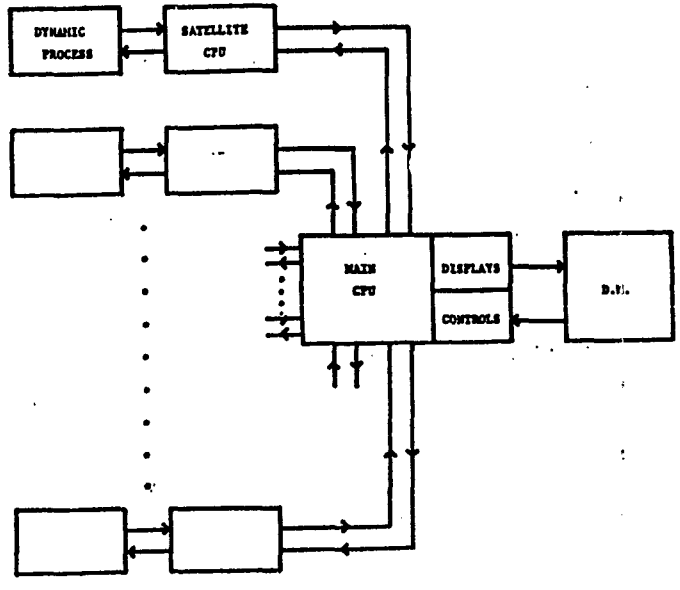


Figure 1. Hierarchial Control of Large Scale Systems with the Human Operator as the Supervisor of a Distributed Computer Network.

In the above figure we can treat the messages sent to the main CPU from a particular satellite CPU as a random noise process with an exponential auto-correlation. The longer the time constants of the auto-correlations, the less frequently the processes need higher-level attention, and the more the number of processes the DM can monitor and control.

We can group the similar processes in 'N' different ensembles. When a task is created by the satellite CPU of a particular process it will queue for the action of the DM in the particular ensemble of the process. Each queue may be characterized with a different mean interarrival time between the tasks; furthermore there may be "transition time" losses $\tau_{i,n}$ for the DM when he transfers his action from the i.th queue to the n.th one. (Ref. 4)

In each queue different tasks (i,j), $i = 1, 2, \ldots\ldots, N$

$j = 1, 2, \ldots\ldots, M_N$

-- where 'i' is the index number of the ensemble to which the task 'j' belongs -- may be created throughout the operation of the system and are worthy of the DM's attention and action.

Each task (i,j) will be characterized with a finite number of variables to indicate:

1) How far away the task is from the 'deadline' for successful action on it -- the 'POSITION' of the task.
2) With what 'SPEED' the task is moving to this deadline.
3) The 'DURATION' of the task.
4) The 'PRODUCTIVITY' of the DM for that task (or group of tasks).
5) The 'VALUE DENSITY' of the task to indicate the benefits accrued per unit time the DM acts on it. Value can then be earned either as the time integral of value density of tasks acted upon, or alternatively, the time integral of the value densities of only such tasks that are successfully completed.

At this point we can observe that the first two factors together specify the amount of time the DM has - TIME AVAILABLE - to act on the task. Similarly the third and fourth variables combine to give the amount of time the DM has to spend - SERVICE TIME - to successfully complete the task.

Note further that among the first and second, and among the third and fourth variables, there is (or may be) an extra degree of freedom. Thus we can let either 'initial position' or 'speed' be constant and also can let either 'initial duration' or 'productivity' related to the particular task be constant.

In Figure 2 we show these task variables explicitly for task (1, 1) which is queueing for the attention and/or action of the DM, along with other tasks present, namely (1, 4), (2, 1), (2, 3),(2, 5), and (3, 2) and (3, 4).

From the above discussion we can infer that the 'state' of the 'system' that the DM is supervising is the vector:

$$X = (x_{1,1,1}, x_{1,1,2}, x_{1,1,3}, x_{1,1,4}, x_{1,1,5}, x_{1,4,1}, \ldots, x_{i,j,k}, \ldots, x_{N,M_N}, \quad 5)$$

where i, j, k represent the queue, the task in the particular queue, and the variables of this task at a particular time, respectively. Note that the first five terms in the above vector define the position, speed, duration, productivity, and value density associated with the first task in the first queue (1, 1) at the particular time.

In Figure 3 we show part of the state space of a task (which itself is a part of the state-space of the supervised system) to illustrate some of the above-mentioned ideas. Note that the two arrowed curves represent two possible

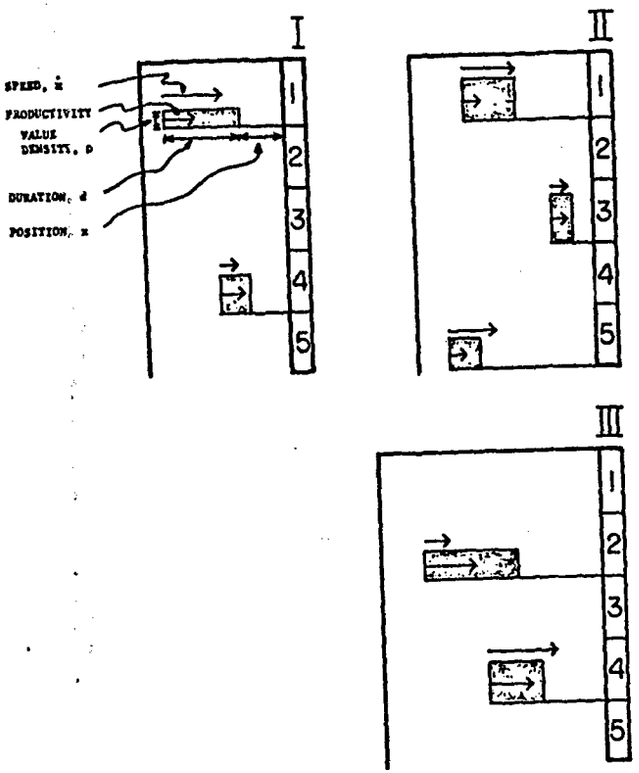


Figure 2. Paradigm of Dynamic Task Demands with Multiple Queues. (Experimental Computer Display)

state trajectories. The successful completion of a task requires that the duration associated with the task go to zero before the position variable reaches zero. The ellipse in the figure represents the boundary of a volume within which is the initial task state at the creation, with 99% likelihood.

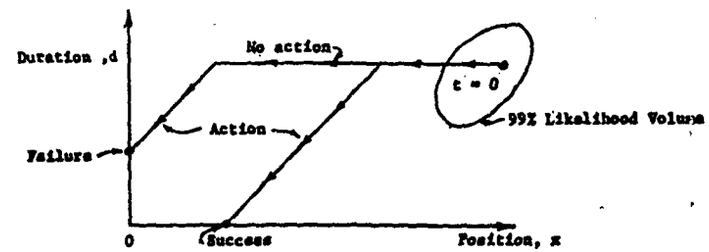


Figure 3. Task State-Space with Two State Trajectories.

We can now model the system as tasks generated according to a probability distribution, and which queue for the DM's action. In particular, we will assume that the tasks are created independently according to a real-time Poisson process with Gaussian distribution of task parameters:

$$P\left[\underline{m}_{i,j+1}(t+\tau), \underline{X}_{oi} \middle| \underline{m}_{i,j}(t)\right] = P(\lambda_i, \tau, \underline{X}_{oi})$$

Note that the parameters of this probability distribution may be different for each queue. The most important parameters in this context are:

1) The mean interarrival times between successive tasks i.e., $1/\lambda_i$
2) The window-length of the queue (W_i), as an indication of the maximum expected position for the tasks on the queue.
3) The mean values and variances of the position, speed, duration, productivity and value density of the created tasks at their initial appearance (and if they do change, their system dynamics).

We are thus modeling the input to the system as an exponentially correlated disturbance. The auto-correlation and the power-spectral density of this process are shown in Figure 4. Note that this input can be considered as the output of a first-order shaping filter with characteristic Φ_{xx} to which white-noise is the input.

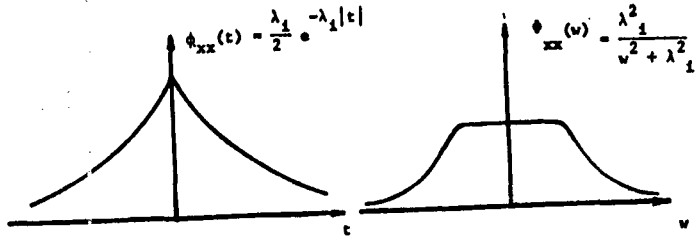


Figure 4. Autocorrelation and Power Spectral Density of Input

The responsibility of the DM is to choose among the alternatives. He will act on -- and therefore control -- the task that he chooses. He usually can act only on one task at a time, although he can time-multiplex his control action by switching from one task to another. This idea is very close to the 'Bang-Bang' control (Ref. 5), where the control action is on the boundary of a feasible control space, with the added hard-constraint that there be only one task that is being acted upon.

OPTIMIZATION

In choosing his control - i.e., which task to act upon - we can model the DM as an optimal controller who maximizes his expected returns over a planning horizon. Dynamic Programming (Ref. 6) is the most promising technique to use in the above mentioned Supervisory Control Paradigm.

In particular, the DM will act to maximize his expected total returns over a finite planning horizon, and perhaps with a non-zero discount function f(t) : --which can for example be $\exp(-\beta t)$--

$$\max E\left[\int_0^T R(t)\ f(t)\ dt\right]$$

where $R(t) = \sum_{ij} R_{ij}(t)$

in which the summation is over all the tasks that the DM expects to act over his planning horizon. $R_{ij}(t)$ is the return he gets for acting (or completing) the task (i,j) during (or at) time t.

For the case in which the DM gets credit continuously while acting on a task, the $R_{ij}(t)$ will be as shown in Figure 5.

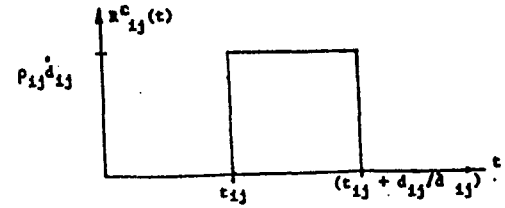


Figure 5. D.M.'s Return Function for "Continuous Credit While Acting".

In this figure, t_{ij}, ρ_{ij}, d_{ij}, $\dot{d}_{ij}$ represent the time at which the DM plans to start acting on the task, the value density of the task, the duration of the task, and the productivity of the DM for the task, respectively.

If however, the DM is going to get credit only after successfully completing a task, then $R_{ij}(t)$ will be as shown in Figure 6.

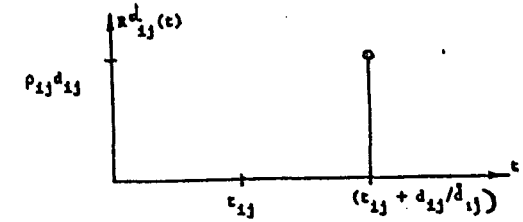


Figure 6. D.M.'s Return Function for "Credit Only When Finished".

The DM will, in effect, at each decision point, choose a number of tasks -- the number being a function of the planning horizon and the mean interarrival times-- in a linked list that he intends to act upon to maximize his expected returns, and then he will actually act upon the first task in this list.

It is probable and acceptable that he might have to give up on acting on some tasks when their 'available times' are small - due to their high speed and/or due to their proximity to the deadline - or when they have comparatively low value densities, especially in competition with other simultaneously available tasks which are preferred in these respects. Another important parameter, of course, is the transfer time between the queues. He has to consider the fact that he will end up getting no credit for a period of time when he transfers his control from the i.th queue to the n.th one. Note that when this Transition-Time matrix is a null matrix, i.e., when there are no transfer time losses and when the DM is continuously being awarded for the task he is serving, then he will do instantaneous maximization, as this will satisfy the maximization of expected returns over the long term too.

The parameters T and β which are the planning horizon and the discount rate will directly affect each other. There also is a physical limit on the planning horizon: it can (or should) not be greater than the time left to the end of the experiment.

$$T < (T)_f - t$$

From the form of the optimization equation above we can see that the discount term might be thought of as inversely proportional to the planning horizon, since the latter is the limit of the integral within which is the exponential decay term of the former. However, there are no physical limits on the discount parameter, other than it being positive semi-definite.

ESTIMATION

Up to now we have talked about the human decision maker as an optimizer. However in many cases estimation of the structure and the state of the system that he is controlling will be an integral part of his optimal dynamic decision making (Ref. 7).

In the Supervisory Control Paradigm, estimation of the state will especially be important when the speeds of the tasks are high, when the transfer time from the present queue to another one is significant, and/or when the DM is not allowed to observe another queue while acting on a task in the present one.

We have noted in the previous section that the DM has a finite planning horizon. We also note that some part of this time-horizon will be (or may be) displayed to the DM, since he will have access to a finite 'window' through which he can observe the tasks that are in the vicinity of the deadline. However in some cases this might not suffice to cover the whole planning horizon. In particular we note that the effective contribution of the DM's observation of a particular queue through the displays to his planning horizon will be:

$$L_i / V_i$$

where L_i and V_i represent the mean position and mean speed of the tasks at their initial appearance (creation) in queue i, respectively. He therefore may still have to predict ahead of what is available to him by in effect extending L_i/V_i to T, where T is defined as the planning horizon of the DM.

We note again that it is entirely possible for the DM to use the effects of only those tasks that are observable by him, in making a dynamic decision as to which task to act upon. A possible exception will be the case of not being able to observe the other queues when acting on the present one.

When making estimates about the non-observable tasks, however, the DM will use his own estimated values of the system variables - like mean position L_i, and mean speed V_i - which are presented to him as random variables, and maybe even τ_{ip}, which, although is presented as a deterministic quantity, might be perceived as a random variable by the DM, perhaps due to his own noise creation. We might add that if the DM is 'trained' in interacting with (controlling) the system, then his subjective probability distributions about the state of the system will not be too much different from the actual ones (Ref. 8).

We can model this estimation process as one in which the DM uses his a priori subjective probability distributions before making judgments at each instant, and then updating these subjective distributions according to what he has observed. This estimator then becomes a recursive filter, i.e., there is no need to store past measurements for the purpose of computing present estimates. We therefore are going to model this updating as one in which the DM will 'learn' about the Gaussian (and Poisson) properties of the queues while he is controlling the system (Ref. 9) - unless he was already trained about them — in a maximum likelihood/minimum variance Bayesian way. The DM will then use these a posteriori subjective probabilities in making his next decision to maximize his expected gains. It is interesting to note here the parallelism between this approach and the Kalman Filter used in Linear-Quadratic-Gaussian control theory, in which it is assumed that the physical properties of the Gaussian random processes are known exactly, and then the future states are estimated based upon the present states and the Gaussian properties (Ref. 10).

Taking the estimation into account, we are now ready to present a block diagram of our modeling effort:

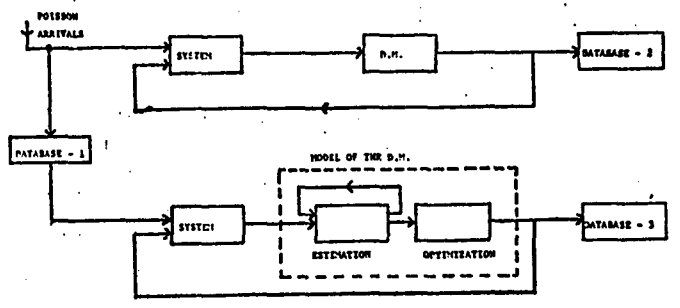


Figure 7. Block Diagram of the Modeling of the Decision-Maker.

EXPERIMENTS & Preliminary Results

Due to the speed and programming restrictions of our Interdata-70 computer/Imlac-display pair we currently have 2 programs for the experimental set-up. The first program is the interactive one. It essentially creates the tasks and displays them to the DM as in the form shown in Figure 2. The DM is able to choose one of these by pressing a pen on the data-tablet. At each iteration the choice of the DM is recorded on the disk. Also recorded in another database are the Poisson arrivals of the tasks and their parameters.

After the experiment we use these two databases to print out the time history of the DM's actions under the given circumstances at that instant. Figure 8 shows a particular subject's control action history of a 3-queue, 5 task per queue (statistically max.) system. Note that the two numbers at the top of each dotted sequence indicate the queue and the task number, respectively. Also at the appearance of each task, two numbers are written in the figure to indicate to the reader the service time the task requires and the time span the task would be available on the screen. The numbers after the stars indicate the relative value densities of the tasks, relative to the average of the mean values of this parameter in each queue.

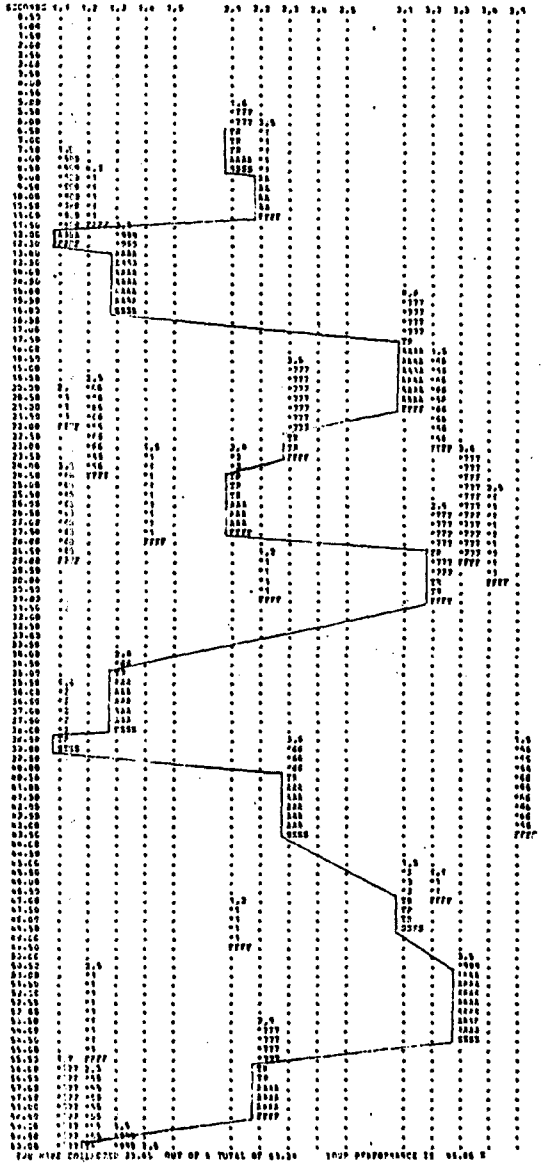


Figure 8. Record of D.M.'s Action Trajectory in Real Time

The A's indicate action by the DM on the particular task at the given time -- written on the left hand side of the Figure. If the task was successfully completed 'SSSS' is written at the time of completion. If, however, the task was not successfully completed before it disappeared from the screen, then 'FFFF' is printed at the time of disappearance.

The second program uses the databases created during the experiments with the subjects, and does estimation and optimization at given points in the parameter space, as shown in Figure 7.

In Figure 9, we show the action trajectory of the optimal model for the partial credit mode as shown in Figure 5. The first thing that attracts attention while comparing Figures 8 and 9 is that the model responds more quickly to new tasks than does the human. To compensate for this we can adjust the model to simulate the human "Response-Time" r_{in} which adds to the τ_{in} 'Transfer Time' loss between queues i and n. This'response-time' loss is due to decision time loss, neuro-muscular lags, and the time losses for physical hand movement on the data-tablet. In our preliminary experiments we have noticed that the effects of this response-time loss on the human's performance to that of the optimal model (i.e., operating at zero response-time loss) decreases rapidly as the 'available-times' of the tasks are increased.

Figure 10 is a comparison of value gainud by the 'Optimal Control' of Figure 9 (operating at $r = \emptyset$, $\beta = \emptyset$, T = time until the end of the experiment) with the control output of a subject as shown in Figure 8. Figure 11 shows the same data for a second subject.

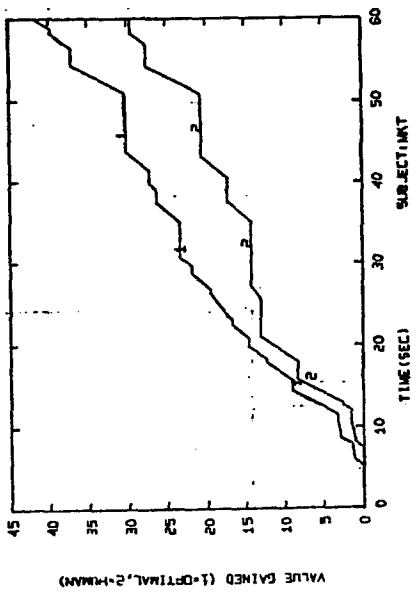


Figure 10. Time Histories of Value Gained by the DM and the Model of Figure 9. (Subject I).

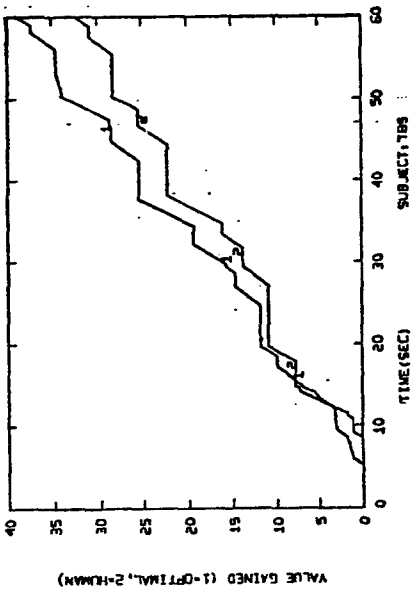
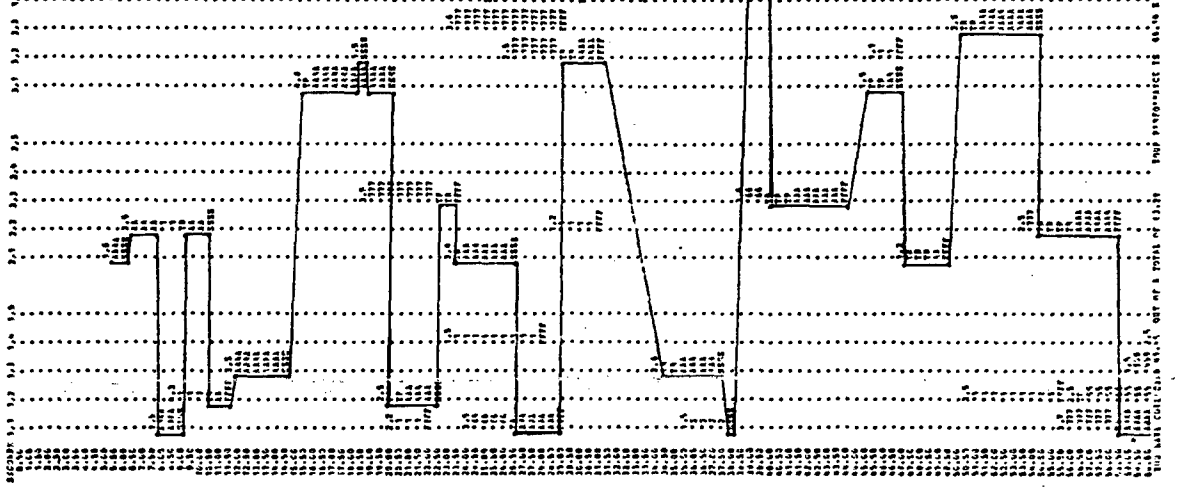


Figure 11. Time Histories of Value Gained by the DM and the Model of Figure 9. (Subject II).

Figure 9. Record of Model's Action Trajectory in Real Time



By using a discount function f(t) for future returns

$$f(t) = 1 - \beta(t),\ 0 < t < \frac{1}{\beta}$$

in the objective function, we can search the parameter β space and find the best β that can be employed to get maximum returns under given task conditions like interarrival times, transfer time delays, mean values and variances of task states, etc. While searching this β-space we also compare the time histories of cumulative value gained by the human vs. the model - as in Figures 10 and 11 - and can identify the discount parameter for the human as the β which minimizes the least squares difference between these two curves. As shown in Figures 12 and 13 the parameter β that best identifies the human is not necessarily the β that maximizes the total value gained at the end of the experiment.

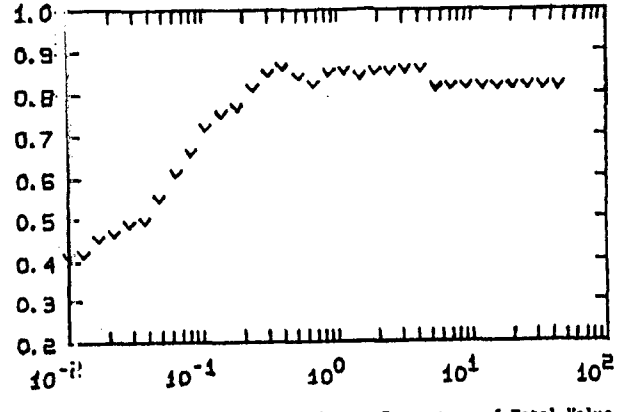


Figure 12. Value Gained by the Model as a Percentage of Total Value Offered with Varying β-discount Rate.

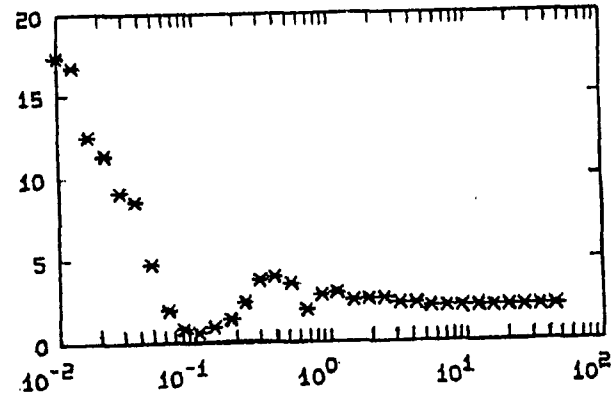


Figure 13. Least-Squares Difference Between the Model and the Human as a Function of β-discount Rate.

We are doing the identification under different criteria as well. These are:

1) minimization of the sum of the squares of the differences in instantaneous or incremental values simultaneously gained by the human and by the model.

2) maximizing the fraction of total experimental time the model and the human are simultaneously serving the same tasks.

3) maximizing the number of tasks which eventually are served by both the model and the human independent of when they are served or for how long.

REFERENCES

1. "The Smart Machine Revolution", Business Week, No. 2439, July 5, 1976.

2. J. C. R. Licklider, "Man-Computer Symbiosis", I.R.E. Trans. on Human Factors in Electronics, V. HFE-1, No. 1, March 1960.

3. E. Palmer, "Pilots Manual for the 4D Area Navigation and the Autopilot Systems in the Flight Management Project Cockpit Simulator", N.A.S.A. Report.

4. T. B. Sheridan, "Optimum Allocation of Personal Presence", I.E.E.E Trans. on Systems Science and Cybernetics, V. SSC-6, No. 3, July 1970.

5. A. E. Bryson and Y-C. Ho., "Applied Optimal Control", John Wiley & Sons, (1975 - revised printing).

6. A. Rapoport, "Dynamic Programming Models for Multistage Decision-Making Tasks", J. of Math. Psychology, V. 4, pp. 48-71, 1967.

7. H. S. Witsenhausen, "Separation of Estimation and Control for Discrete Time Systems", Proceed of the I.E.E.E., V. 59, No. 11, November 1971.

8. K. S. Fu "Learning Control Systems - Review and Outlook", I.E.E.E. Trans. on Automatic Control, V. AC-15, No. 2, April 1970.

9. D. G. Kenhn, "A Note on Learning for Gaussian Properties", I.E.E.E. Trans. on Information Theory, V. IT-11, No. 1, January 1965.

10. A. Gelb, ed., "Applied Optimal Estimation", M.I.T. Press, (1974).

N79-17495

A MODEL OF THE HUMAN SUPERVISOR.

Jan J. Kok , Ron A. van Wijk ').

Man-Machine Systems Group, Laboratory for Measurement and Control, Department of Mechanical Engineering, Delft University of Technology, Delft, Netherlands.

0. Summary.

The outline of a possible construction of a general model of the human supervisor's behavior is given. In the description of the model very much attention is paid to its basic philosophy and the underlying thoughts and concepts - the system theoretic evaluation is only briefly considered at the end, where some results are given of an application of the model. Based on the notion of the state of the supervised system, which contains all information needed for the examination of all system functions as well as the deduction of adequate correction signals, three basic submechanisms are accepted to constitute the model - the observer/reconstructor part, the decision-making part, and the controller part. For the relations between the task variables (system dynamics, noise parameters, supervisory task) and the parameters of the different submechanisms of the model a set of hypotheses is postulated. Then the verification of the model hypotheses is considered using variations in the task variables. For that purpose the three basic submechanisms are given in terms of the operational structures. For the observer/reconstructor part an optimal observer is introduced, for the decision-making part the description is given in terms of decision rules, and the controller part is developed along the line of a minimum-time criterion.
Finally, for the identification of the model parameters an approach is suggested which makes use of a multi-dimensional error criterion. Each of the elements of this multi-dimensional criterion corresponds to a certain aspect of the supervisor's behavior, and is directly related to a particular part of the model and its parameters. This approach offers good possibilities for an efficient parameter adjustment procedure.

1. Introduction.

The ever-increasing automation in industry has put the human operator more and more in a new role relative to the machine - the role of a supervisor over an automatically controlled system. This means that at the occurence of deviations, failures and disturbances in the operation of the system the human supervisor has to perform a task of corrective actions, i.e. he has to detect the errors and to decide if corrections are necessary. Therefore, the supervisory task involves monitoring and observation as well as decision-making and control.

For the design of adequate man-machine interfaces a good knowledge and understanding of the supervisor's behavior is necessary. Descriptive and normative models of the supervisor can add significantly to the, still missing, know-how. However, to be useful in design of automated systems of diverging complexity and character, the new models should be valid over a broad range. This calls for models based on a general framework, which can be applied to a variety of situations. Empirical models, developed for a particular set-up, are of no interest.

') This research is partially supported by the Netherlands Organisation for Advancement of Pure Research (ZWO).

In this paper we give an outline of a possible construction of a general model of the human supervisor's behavior. We are very aware of the fact that our attempt is just a first step towards the final goal of a useful tool in the design of man-machine systems, for instance with respect to the limited number of aspects of the supervisory task which is taken into account. In the description of the model we pay very much attention to its basic philosophy and the underlying thoughts and concepts, leaving the system theoretic evaluation to the last section where we give some results. The reason for this approach lies in our wish narrow the gap between experimental psychology and system engineering modelling, which, in our view, would be very desirable.

2. Evaluation of a supervisory model.

The human supervisory task can be considered as a control task, but on a higher hierarchical level, due to the fact that the direct control of the system is accomplished by the automatic controller. Therefore, a control model can serve as a starting point in the evaluation of a supervisory model. For this purpose, in our view, the optimal control model (Baron, Kleinman, Levison [1]) is the best choice, because of the following reasons:

- composed of some basic submechanisms it is a normative model which on the one side allows the system theoretical effectuation, and on the other side presents possibilities for interpretation of the parameters from psychological and behavioral sciences point of view.
- essentially this model is capable of handling very complex situations as encountered in supervisory control tasks.

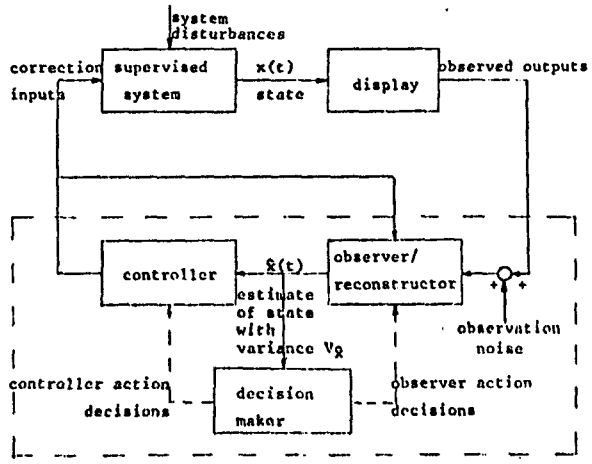


Figure 1. Basic concepts of the supervisory model.

As in the optimal control model the basic underlying thought of the model of the human supervisor given here is the separation of the model in a observer/reconstructor part and a controller/decision making part. This separation is based on the concept of the state of the supervised system, for the state contains all the information needed for the examination of all system functions as well as the deduction of adequate correction signals. Since the system state itself is not directly available, from all the information at the supervisor's disposal, i.e. the observed system outputs and the correction inputs, and the knowledge of the system dynamics (the internal model of the supervisor), at the best an estimate or reconstruction of the state can be obtained: this estimate is generated by the observer/reconstructor part of the model. The estimate of the system state obtained this way and the statistical uncertainties related to this estimate, are the sources of information upon which the controller/decision making part of the model can come to corrections and intervening actions. The control strategy and the decision rules relevant to these corrections and actions follow from the system dynamics and the supervisory task to be performed by the operator. The basic concepts of the supervisory model, i.e. the separation in three submechanisms, is given in Fig 1.

In the following sections we will discuss the assumptions and the hypotheses upon which the evaluation of each of the submechanisms can be based. For reading convenience, however, we first give some definitions of terms and quantities involved in the modelling process.

3. *Definition of terms.*

Supervised system: the system to be controlled and monitored, or, in general, to be supervised. This includes the process as well as the automatic controller.

System disturbances: noisy inputs to the supervised system which cannot be controlled by the supervisor.

State x(t) of the supervised system: collection of (physical) quantities in the system which, expressed in a momentary value, contain all information on the affectuation of the system inputs in the past.

Display: system which provides the representation of the state x(t) of the supervised system to the supervisor; usually, a reduction of the explicit information content of x(t) takes place here.

Observed outputs: the signals giving information about the state of the supervised system which are available to the supervisor, and which are obtained by observation of the display.

Observation noise: noisy signal which represents the supervisor's uncertainty about the observed outputs.

Observer/reconstructor part of the model: (dynamic) system which generates a reconstruction of the state of the supervised system from the information received from the observed outputs and the correction inputs. Due to the uncertainties in the available information (as a result of the observation noise and the unknown system disturbances), and the reduction of the information by the display, this reconstruction cannot be better than an estimate with a certain error.

Estimate $\hat{x}(t)$ of the state: reconstruction of the state of the supervised system, which takes into account the statistical properties of the uncertainties in the available information. The estimation error is characterized by the estimation variance $V_{\tilde{x}}$.

Controller part of the model: a system that generates a correction signal for the supervised system from the obtained estimate of its state (usually in the form of set points for the automatic controller).

Decision making part of the model: a system that, based on the estimate of the state of the supervised system and its error variance, comes to decisions about the realisation of control actions (the adjustment of set points) and the observation actions (the sampling of a controlled variable, for instance).

Correction inputs: the signals generated by the supervisor to make corrections to the controlled variables of the supervised system.

4. Model hypotheses.

For the relations between the task variables (sytem dynamics, noise parameters, supervisory task) on the one hand and the parameters of the different submechanisms of the model on the other hand, the following hypotheses are essential:

- The observer/reconstructor part of the model is assumed to be independent of the supervisory task to be performed by the supervisor, and thus only depends on the supervised system (to which system do the observed outputs belong, and, therefore, which interrelation between them does exist?), the display (in which way are the state variables represented in the observed outputs?) and the noise parameters (what are the characteristics of the uncertainties in the observed outputs?). Along system theoretical lines the concept of the observer/reconstructor part can be affectuated for a given set of system parameters and noise parameters. The result shows the reflection of the supervised system and the display within the observer/reconstructor part of the model. This fact offers the oppertunity to the introduction of an internal model concept to model the supervisor's knowledge of the system dynamics. Assuming a complete notion of the dynamics of the system by the supervisor, the statistical properties of the observation noise are the parameters of the observer/reconstructor part of the model.
- The controller part of the model is assumed to be independent of the observation/reconstruction process, and thus only depends on the supervised system (which system will receive the correction signals, so in what way will these signals influence the state of the system) and the supervisory task (what are the controlled variables, and which control effort is allowed for the necessary corrections?). Also for this control problem the solution is provided by system theoretical elaboration, which offers several possibilities to model the supervisory task (for example, is the supervisory task given in terms of a quadratic performance index, or in terms of a minimum time criterion?). Once more, the internal model concept can be introduced to model the subjective interpretation of the human supervisor of the supervisory task laid upon him. This is an extension of the internal model concept mentioned before. The weighting factors in the supervisory task are the parameters of the controller part of the model.

- The decision-making part of the model is assumed to operate on the controller part as well as the observer/reconstructor part of the model. This is because of the fact that the human supervisor must make decisions to come to corrective control actions, but also should watch the observation process and, if necessary, take action to avoid inadmissable uncertainties in the estimates of the state of the supervised system (this kind of decisions arises, for instance, when the supervisor has the opportunity to sample from time to time the controlled variables more accurately than can be deduced from direct observation of the continuously available outputs). Therefore, the total decision-making process can be separated into two different groups of decisions, viz. the decisions related to control actions and the decision related to observation actions. A general concept which can be used as the basis of the decision-making part of the model, and which postulates the relation between task variables (the system dynamics, the display, the supervisory task, the noise parameters) and the decisions or the decision rules, is not available yet. To be able to develop such a concept in the future, the model of the decision-making part will be based, temporarily, on a set of free parameters, given in terms of decision rules and the crossing of certain threshold values. After determination of these free parameters, i.e. the threshold values, in different experimental situations, a concept could be developed for the modelling of a relation between the threshold values and the independent task variables. With regard to the decisions concerning control actions, the threshold values are related to the correction signals (a control action will take place if the necessary correction exceeds the particular threshold value), and for the decisions concerning observation actions the threshold values are related to the uncertainties of the estimates of controlled variables which can be sampled (a sample will be taken if the uncertainty of the estimate of the controlled variable exceeds the particular threshold value). It should be noticed that the information needed for these decisions in terms of the correction signals and the uncertainties of the estimates can be directly derived from the estimate $\hat{x}(t)$ of the state and the error variance $V_{\tilde{x}}$, making use of the control algorithm and the parameters of the display, respectively. The threshold values in the decision rules are the parameters of the decision making part of the model. In fact, these parameters are *not* independent quantities, but are related to the supervisory task, the noise parameters, and the dynamics of the supervised system and the display. This relation will be considered as a topic for future research.

The outline of the modelling concept introduced above has a very general character. The model gives a qualitative as well as a quantitative description of the human supervisory behavior, making it interesting from psychological and system theoretical point of view. The system theory provides the necessary evaluation of the submechanisms of the model, while the parameters involved are problem related quantities which have a direct meaning, and practical significance (human uncertainties in the observation process, the supervisory task, the system dynamics, the thresholds of the decision rules).

The fundamental separation between the observer/reconstructor part and the controller/decision-making part of the model is an attractive feature of the postulated model, which allows the experimentator to choose the variation of task variables appropriately, and so study each of the resulting effects individually. The verification of the model hypotheses will be considered in the following section.

It should be noticed that in the description of the model given above it was assumed that the model will be applied to supervisory control situations where the time constants of the system exceed values of 10 seconds and more, and thus that the human control capacities are not severely limited by his inherent reaction time or his muscular system (in the applications of human control of fast response systems modelled by an equivalent delay time).

A final remark concerns the modelling of the human uncertainties in the perception of the correction signals. Here two different conceptual solutions could be suggested. According to the first concept, which is commonly accepted as, for example, in the optimal control model (Baron, Kleinman, Levison [1]), these uncertainties can be modelled by a motor noise which is additive to the correction (control) signal; this approach was shown to be successful in the case of modelling the human control behavior, where this noise in fact represents the uncertainties of the operator's control movements. The second concept, however, might be more appropriate for the supervisory control situation. Rather than by a motor noise, the human perceptual uncertainties are now modelled by an additional observation noise, which acts on the feedback signal to the observer/reconstructor part of the model only. Which one of these approaches gives the best description of the human supervisory behavior can be determined by direct application of the two concepts on practical data, which will be part of the investigation.

5. *Verification of the modal hypotheses.*

Explicit verification of each of the model hypotheses is very difficult, if not impossible, due to the strong interaction of the model submechanisms. What can be done, however, is the verification of the influences on the model of each of the task variables, which implicitely verifies the total set of hypotheses. The following variations of task variables are to be considered:

- *Variation of the intensity of the system disturbances.*
  Variations of the intensity of the system disturbances will conceptually influence the observer/reconstructor part of the model, but not the controller part. Assuming an optimal full-order observer as the observer/reconstructor part of the model, the relation between the system noise intensity matrix and the observer gain matrix as well as the error covariance matrix is completely prescribed by the system theoretic solution of the stochastic optimal reconstruction problem. An increase of the system noise intensity, relative to the other noise sources in the closed loop, will also result in an increase of the observer gain matrix, which reflects the fact that more attention should be paid to the observed outputs (increasing the speed of the state reconstruction). Also, an increase of the resulting error variance matrix, which corresponds to an increase of the estimation uncertainty, should coincide with an increase of the repetition rate of the observation actions of the decision-making part of the model. However, the submechanisms of the controller part and the decision-making part of the model should be invariant under the variations of the intensity of the system disturbances.

- *Variation of the display.*
The display is reflected in the observer/reconstructor part of the model, so variation of the display parameters should be effective here. As far as the variations affect the observability of the supervised system (which information is available in the observed outputs for the reconstruction of the state), the resulting error variance will be affected too. Again, such variations and the resulting error variance and observer gain matrix are prescribed by the system theoretical solution of the stochastic optimal reconstruction problem, and can be verified accordingly. Additional to these consequences of variations of the display parameters also a drift in the observation noise characteristics may be expected, which might change the whole picture. This must be examined very carefully. The submechanisms of the controller part and the decision-making part of the model should remain the same for various displays, however.
- *Variation of the supervisory task.*
Variations of the supervisory task are conceptually of no influence on the observer/reconstructor part of the model, but are only of importance to the controller part of the model and probably also to the decision-making part (i.e. the threshold values). The supervisory task can be varied, for instance, by changing the admissable control effort by proper instruction of the supervisor. Depending on the assumed structure of the supervisory task, such as a quadratic performance index or a minimum time criterion with limited signal magnitudes, either a relative weighting of control effort against result will take place, or an actual limitation of the admissible corrections. The choice for the modelling of the supervisory task also determines the nature of the resulting correction signals. In case of a quadratic performance index the model will generate a continuous correction signal, whereas in case of a minimum time criterion a discrete signal will result (bang-bang character, e.g. Crossley and Porter [2]). Noting the fact that in supervisory control situations the human supervisor generates signals that are rather discrete than analogue, it is very likely that a model based on the minimum time criterion will give better results as far as the input-output description concerns of the supervisor's behavior. This expectation was confirmed by application of such a model on practical data (Hupkes [3], Beliën [4]). As the parameters in the minimum time criterion, the magnitude of the admissable corrections can be considered or the length of the time interval during which a perceived deviation of the controlled variable must be corrected by a constant input to the system. The relation between these parameters is determined by the dynamics of the supervised system.
For the relation between the supervisory task and the threshold values of the decision-making part of the model, no modelling concept has been postulated yet. In our investigation we follow the approach of determining first the values of the parameters of the minimum time criterion and the threshold values for different supervisory control situations, which give an optimal description of the human input-output relation. After that, we intend to analyse the results and come up with a modelling concept which will fit the data, and which gives the threshold values in terms of the supervisory task and the dynamics of the supervised system.

- *Variation of the dynamics of the supervised system.*
The dynamics of the supervised system are involved in all mechanisms of the model, thus in the observer/reconstructor part as well as the controller and decision-making part of the model. As far as the influences on the observer/reconstructor part and on the controller part are concerned the postulated modelling concepts can be evaluated and verified by variation of the dynamics of the supervised system. For the influences on the decision-making part this working scheme must be modified, since there was no causal relation postulated between the system dynamics and the threshold values. Here, a similar procedure must be followed as was suggested for the investigation of the relation between the threshold values and the supervisory task. And again, the final objective will be to develop a modelling concept which gives the link between these parameters.
Careful attention should be paid to the influence of the dynamics of the supervised system on the observer/reconstructor. The structure of this part of the model is such that it includes a reflection of the supervised system and the display, which can be considered as the internal model of the supervisor. It is not very likely that, regardless of the complexity of the supervised system, the internal model can be assumed to be an exact copy of this system and the display. This question will be subject to close examination by varying the order of the supervised system and matching full-order observers as well as reduced order observers in the observer/reconstructor part of the model. Also lower order simplified models could be substituted, giving the opportunity to compare the description of the supervisor's input-output behavior obtained in this way with other models. So far, we have no experience with these techniques yet, due to the fact that our experimental set-up involves a low order supervised system which is too simple to raise this problem.

6. *Identification of the model parameters.*

Before attention can be paid to the verification of the model hypotheses, the crucial problem must be considered of how to compare the model behavior with the behavior of the supervisor. The most important quantitative measure for the comparison is given by the resulting correction signals generated by the model and the human supervisor. The most common technique makes use of a criterion which is defined on the error of the two signals, usually in the form of the time average of the squared error. The parameters of the models are then adjusted such that this criterion has a minimum value. This method is quite satisfactory in most cases if continuous signals are considered. In our case, however, we are dealing with signals which are rather discrete varying in time than continuous. Application of the quadratic error criterion on these discrete signals, where only the amplitude of the signal is taken into account, leads to bad results (Hupkes[3], Beliën[4]). To achieve a realistic mutual resemblance of the discrete signals it turned out to be very essential to define the error criterion not on the amplitude only, but also on the time instants of changes in the amplitude of the correction signal.
A similar problem arises in the comparison of the other quantitative output of the model and the human supervisor, viz. the series of observation actions. Here only the time instants are of concern, and the amplitudes are of no importance.

As a result of the considerations given above, we would like to suggest an approach to the problem of comparison of the model and the human supervisor, which is based on the following criterion elements:

- the instants in time where the observer actions are taken (related to the relevant submechanism of the decision-making part of the model),
- the instants in time where controller actions are taken (also related to the decision-making part of the model, but now to the other submechanism),
- the amplitudes of the correction signals (related to the controller part of the model).

This approach has the advantage of a direct coupling of the relevant parts of the model with each of the elements of the criterion. Also, along this line, the individual consideration becomes possible of the different aspects of the resemblance of the supervisor and the model. A consequence of the distinction between the different aspects is the introduction of a vector valued criterion function (e.g. Steward, Kavanaugh, Brocker [5]), which complicates the parameter optimization procedure. On the other hand, the isolation of these different aspects and the direct relation with the related parts of the model offers the possibility of a successive optimization of each of the parameters involved, which can be performed outside the loop. Assuming a given model of the observer/reconstructor part of the model, the threshold values of the decision-making part relevant to the observer actions can be adjusted to achieve a best fit of the measured series of instants. Then, having available the estimate of the state of the supervised system generated by the combination of the observer/reconstructor part of the model and the optimized series of observation actions, the procedure can be repeated with regard to the instants of controller actions and the other threshold values. Finally, the best fit of the amplitudes of the correction signals can be obtained by adjustment of the parameters in the controller part. If necessary, the procedure can be iterated for the adjustment of the parameters of the observer/reconstructor part of the model (see Fig. 2).

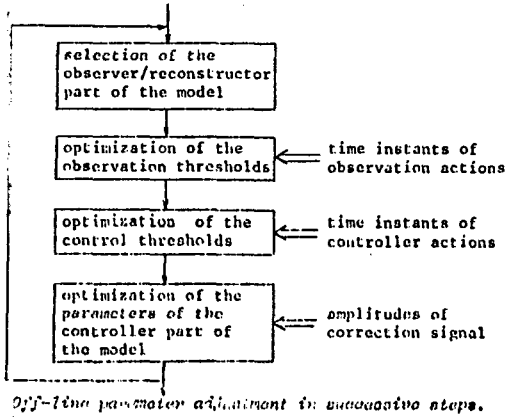


Figure 2. Off-line parameter adjustment in successive steps.

The successive parameter optimization procedure, as well as the combined optimization of the total set of parameters in the vector valued criterion function are now being tested. For the optimization of the vector valued criterion function we make use of a random search method (Steward, Kavanaugh, Brocker [5]), which avoids some complications as encountered in the application of gradient techniques, like termination of the procedure in local minima and the conceptual problem of the non-existence of the derivatives for some parameters.

7. *An application of the modelling concept.*

The model discussed before has been applied to a simple supervisory task in a semi-automatic control situation (see Fig. 3). For a given system, which is disturbed by noise, the quality output $y_1(t)$ must be supervised by the human operator. The momentary value of this quantity cannot be observed continuously by the supervisor, but on his request, will be sampled and become available to him. However, these samples are delayed in time due to processing analysis. The other system output $y_2(t)$ is the feedback signal for the automatic controller, and is also observed by the supervisor. The output $y_2(t)$ is related to the quality $y_1(t)$, but is more corrupted by noise. Using the available information, the supervisor must make proper adjustments to the set points of the automatic controller in order to maintain a good quality. Also, taking into consideration the costs of sampling, he must decide whether or not to ask for a new sample to obtain more accurate, but delayed, information.

In the experimental set-up a digital computer is used to simulate the system and to generate the signals. The subject sits behind a panel, where the information is displayed and the set points can be adjusted. To avoid boredom, two systems of the same character are to be supervised simultaneously by the subject. For this situation, which is commonly found in industry,

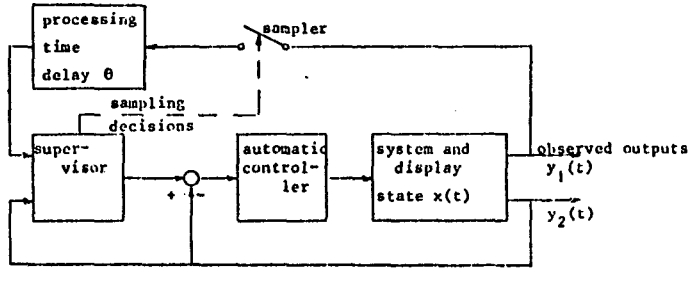


Figure 3. Experimental supervisory situation.

for instance in chemical plants, a model is postulated for the description of the supervisor's behavior, which is based on the concepts given before:
- Using the information about $y_2(t)$ and the last sample $y_1(t_k)$ the supervisor will make an estimate $\hat{x}(t)$ of the state $x(t)$ of the system. The variance of the estimation error of this estimate will be determined by the uncertainties in the observed quantities and by the time elapsed since the last sampling instant.
- Depending on the supervisory task to be performed, the value of the estimate of the state, and the uncertainty of this estimate, the supervisor will make two kinds of decisions:
  a) Decisions to ask for a new sample of the quality signal, taking into account the costs of sampling (observation action decision).
  b) Decisions to adjust a new set point for the automatic controller (controller action decision).
- If the supervisor has decided to correct the set point of the controller, he will select a new value using his information on the estimate $\hat{x}(t)$ of the state of the system.

The system theoretic evaluation of the different parts of the model will not be given in detail. We only give some general comments.

The observer/reconstructor part of the model consists of an optimal full-order observer (Kalman filter) acting on the continuously available output $y_2(t)$. The sampled observation $y_1(t_k)$, which becomes available at time $t_k+\theta$, where θ is the processing time delay, is assimilated in the filter by setting it back in time, process the additional (low noise) information resulting in a better estimate $\hat{x}(t_k)$, and, once again, use $y_2(t)$ in the interval $(t_k, t_k+\theta)$ to arrive at a better estimate $\hat{x}(t_k+\theta)$. Then, until a new sample becomes available, the filter continues its action using $y_2(t)$. The result is an optimal estimate $\hat{x}(t)$ of the state of the system, which is continuously available along with its error variance. From this estimate $\hat{x}(t)$ of the state also an estimate $\hat{y}_1(t)$ of the quality output, and its error variance, are derived using the system output equation.

The decision-making part of the model was designed according to the general outline given in section 4. This means that a decision rule was introduced based on the estimate $\hat{y}_1(t)$ and its error variance resulting in the decisions to ask for a new sample, and another decision rule based on the new set point value to make the necessary correction on the quality output resulting in the decisions to change the set points. The threshold values in the decision rules are the parameters of the decision-making part of the model.

For the controller part of the model we made use of a control strategy which is based on the correction of the perceived deviation of the quality output of the system by a new set point value which is held constant over a certain (minimum) period of time. The length of this time period is the parameter of the controller part of the model.

For the particular experimental set-up the dimension of the state $x(t)$ is three, the outputs $y_1(t)$ and $y_2(t)$ and the set point signal are one-dimensional, and the system disturbances are two-dimensional. So the model includes only three parameters (two threshold parameters in the decision-making part, and one interval length parameter in the controller part), for there are no parameters involved in the observer/reconstructor part, being almost completely determined by the given system parameters, the noise characteristics of the system disturbances, and the prescribed optimal observer algorithms. The only undetermined parameters of the observer/reconstructor part are the intensities of the observation noises in $y_1(t)$ and $y_2(t)$, the first of which is taken to be zero and the other is assigned a value proportional to the variance of the observed quantity itself $(0{,}01 v_{y2}$, according to the experimental data of Baron, Kleinman, Levison [1]). No motor noise has been included in the model, so no parameters are needed to characterize this noise.

The principle of the adjustment of the parameters to match the model outputs with the supervisor's outputs was described in section 6. A three dimensional vector valued criterion, defined on the set point values and the moments of decision to take observer actions and controller actions, was optimized using a random search method. Some results on the outputs of the model are given in Fig.4, which shows that a good fit of the supervisor's behavior can be achieved. Similar results couldn't be obtained using a one-dimensional optimization criterion for the parameter adjustment. Also, testing a model based on the acceptation of a quadratic performance index for the modelling of the supervisory task, the resulting set point variations showed such a complete different character that no good fit of the data could be obtained (Hupkes [3], Baliën [4]).

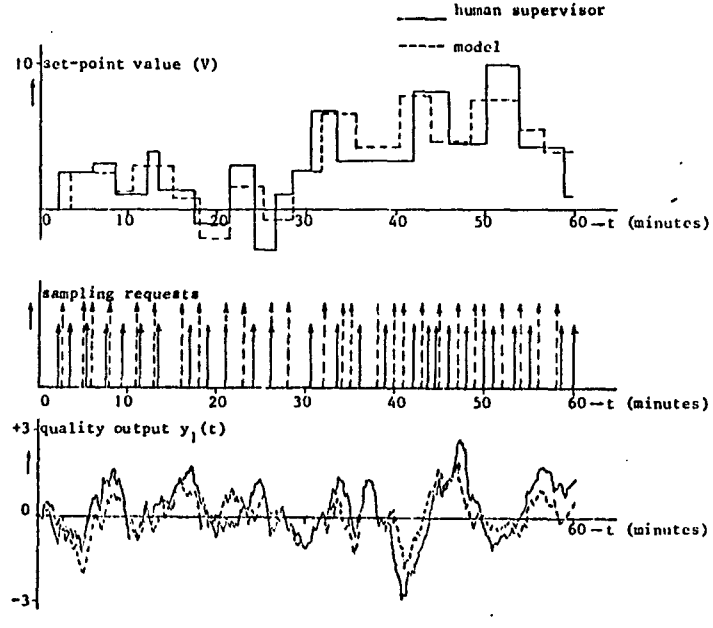


Figure 4. Outputs of the model and the human supervisor.

8. *References.*

1 Kleinman, D.L.; Baron, S.; Levison, W.H.: A control Theoretic Approach to Manned-Vehicle System Analysis. IEEE Trans. on A.C. Vol. AC-16 No. 6 (1971) pp. 824-832.

2 Crossley, T.R.; Porter, B.: Dead-Beat Control of Sampled-Data Systems with Bounded Input. Int. J. Contr. 1974 19,5 (869-876).

3 Hupkes, W.: Het modelleren van het regel- en beslisgedrag van de mens in half geautomatiseerde systemen. Ir-thesis: Delft januari 1977, 55 p. A 212.

4 Beliën, W.L.H.: De realisatie van de proefopzet voor het onderzoek naar het regel- en beslisgedrag van de mens in half geautomatiseerde systemen en de evaluatie van het opgestelde model. Ir. thesis: Delft januari 1977, 51 p. A-225.

5 Steward, E.C.; Kavanaugh, W.P.; Brocker, D.H.: Study of A Global Search Algorithm for Optimal Control. 5th International analogue computation meetings 1968. pp. 207-230.

6 Edwards, E.; Lees, F.P.: The Human Operator in Process Control (book) Taylor & Francis Ltd. London 1974.

THE HUMAN AS A DETECTOR OF CHANGES IN VARIANCE AND BANDWIDTH

Renwick E. Curry
NASA Ames Research Center
Moffett Field, CA

T. Govindaraj
Coordinated Science Laboratory
University of Illinois
Urbana, IL

Abstract

The humans's function in the control of processes is shifting toward a supervisory or monitoring role with the advent of increasing automation. One of the functions of a monitor is the detection of changes or failures in the characteristics of the random process. In this paper we consider the detection of changes in random process variance and bandwidth. Psychophysical thresholds for these two parameters were determined using an adaptive staircase technique for second order random processes at two nominal periods (1 and 3 seconds) and damping ratios (0.2 and 0.707). Thresholds for bandwidth changes were approximately 9% of nominal except for the (3sec,0.2) process which yielded thresholds of 12%. Variance thresholds averaged 17% of nominal except for the (3sec,0.2) process in which they were 32%. Detection times for suprathreshold changes in the parameters may be roughly described by the changes in RMS velocity of the process. A more complex model is presented which consists of a Kalman Filter designed for the nominal process using velocity as the input, and a modified Wald sequential test for changes in the variance of the residual. The model predictions agree moderately well with the experimental data. Models using heuristics, e.g. level crossing counters, were also examined and were found to be descriptive but do not afford the unification of the Kalman Filter/sequential test model which has previously been used for changes in mean.

Sponsored by NASA Grant NGR 22-009-733 from the Man-Vehicle Research Division, NASA Ames Research Center

Introduction

The humans's function in the control of processes is shifting toward a supervisory and monitoring role with the advent of increasing automation. Monitoring seems to imply at least the two functions of performance assessment and failure detection (or change detection), although there is no clear cut definition of monitoring per se (Curry and Gai, 1976, Kleinman and Curry, 1977).

Developing descriptive models of human failure detection requires appropriate blending of theory and experiment. Without data, a reasonable starting point is a normative model such as suggested by Phatak and Kleinman (1972). They propose a model of failure detection and identification which requires a Kalman Filter for each alternate hypothesis. Levison (1971) has used estimation theory as a basis for a performance assessment model with good predictions of experimental data for detecting signal limit exceedances.

The emphasis of this paper is to report on experiments yielding information regarding some of the human capabilities and limitations in detecting random process changes, i.e., the psychophysics of random processes. Previous work has involved detection of changes in mean by human observers (Gai and Curry,1976); in this paper we report on other aspects of random process changes which have important applications in the operational setting, viz., changes in bandwidth and variance. A more complete description of the research is contained in Govindaraj (1976).

Description of The Experiments

General

Experiments were conducted to determine the thresholds for changes in the variance and bandwidth of a random process, and to study the detection behavior for suprathreshold stimuli. The process consisted of the output of a second order shaping filter with the transfer function

$$\frac{K}{(s/\omega_n)^2 + 2\zeta(s/\omega_n) + 1} \qquad (1)$$

and zero mean white Gaussian noise as input. In the above equation K is the filter gain, ω_n is the natural frequency, and ζ is the damping ratio. The output of the shaping filter was displayed on the graphics display terminal as a horizontal line moving up and down inside a grid. All the three parameters, the natural frequency, the damping ratio, and gain K may be changed, but only variations in the natural frequency (the bandwidth) and the variance of the input noise (equivalently the output variance) were considered for the present study of failure detection.

D.2.1

Equipment

A PDP 11/34 computer with graphics capability was used to carry out the experiment. White gaussian noise was digitally generated and used as input to the state equation of (1) which was updated once every 10.5 milliseconds. The subject was seated about 76 cm in front of the screen, the screen being at normal eye level; the standard deviation of the nominal process was 2 cm subtending an angle of 0.0267 radians. The subject held a small box with two switches to indicate his response. A 12 inch diagonal P31 fast phosphor cathode ray tube was used for all displays.

Procedure

Fifteen people participated for approximately four weeks in the threshold experiment. Their background and age ranged from college sophomores, to graduate students in science and engineering, to a senior citizen. For each subject, one session lasted for approximately an hour with no more than one session per day. Two series of experiments were conducted during a session: one for a change in frequency and the other for a change in variance.

The general nature of the experiment was described to each subject on his first day. A brief explanation of the observed process was given in terms of the physical analog of a spring-mass system. The subjects were told that there would not be any definite pattern, since the excitation was random, and that they could only form an idea of 'how far on either side the line moves away from the centerline', and 'how fast or slow it is moving'. They were told to observe the 'average behavior of the line'.

After the procedure was explained, the nominal mode was shown for two minutes and then large parameter changes of both sign were shown to familiarize the subject with the nature of failures. For trials in which changes took place the process started with the nominal parameter values and a failure occurred between 8 and 12 seconds after starting. Though it was obvious, the subjects were told about the nature of these first failures (i.e., increase or decrease) during this phase. One nominal and four failures were sufficient for all subjects to become familiar with the changes.

The observers were required to indicate the type of failure (parameter increase or decrease) on each trial and feedback was presented on the CRT. If the failure was detected and identified correctly, it was a 'correct' response. If any switch was pressed prior to an actual failure, it was a 'false alarm'. If the identification was incorrect, it was 'wrong'. Finally, if the failure was not detected within the available time, it was a 'no detection'. ('Wrong' and 'no detection' are considered 'misses' later on.) After every trial, the result was displayed to the subject on the otherwise blank screen for two seconds. After a blanking period of three seconds, the next trial followed in a similar manner.

Stimuli

Stimulus values were chosen according to the following relation:

$$\frac{P}{P_n} = \exp(S \ln R) \qquad (2)$$

where

P_n - Nominal values of the parameter
P - Changed or failed value
R - Ratio of initial change (R = 10)
S - Stimulus ($|S| < 1.0$).

Threshold Experiments

Threshold experiments were conducted using a modified staircase technique which allowed subject familiarization in the early trials of a session and threshold determination in the later trials. This was accomplished by starting with a large stimulus value and regulating the average rate of decrease of the stimulus magnitude so that 20 or 30 trials were encountered before the stimuli were near threshold.

A set of four nominal processes obtained by the factorial combination of $T = 2\pi/\omega_n = (1.0, 3.0)$ seconds and $\zeta = (0.2, 0.707)$ were used. During any one session, only one nominal was presented. The subjects who participated in the threshold experiments were given all four sets of nominal processes via a balanced Latin square design.

Detection Time Experiment

A second series of experiments was conducted to determine the time taken for detection of a failure as a function of the stimulus level. The general set-up for this series was the same as before. However, the criterion by which the subjects responded was different. It was made clear that the objective was to determine how quickly one detects a failure. The subject was specifically told that he was "expected to detect the failure as quickly as possible without making too many mistakes". Another important difference was in the set of stimuli chosen. Since frequency and variance thresholds were available for the four nominals from the previous experiments, four stimulus magnitudes were chosen with the smallest being slightly higher than the threshold. Four parameter increases and four decreases were mixed to form a group of eight stimuli, and a stimulus was presented from this group at random without replacement.

Results

Threshold Values

An exponential approximation was used to fit the stimulus-trial data. The assumed curve had the form

$$|S_n| = a + \exp(-c*n) \qquad (3)$$

where S_n is the value of the stimulus at the trial number t, and a is the threshold. A least squares fit was found using as initial values the threshold values calculated at the end of each session by taking the mean of the last six peaks and valleys of the stimulus vs trial history. In most cases, these first approximation values agreed very closely with the values found by the exponential fit. The thresholds determined in this manner and averaged over subjects are displayed in Table 1.

Threshold magnitudes were compared for parameter increases vs. parameter decreases using a modified t-test. The null hypothesis that the thresholds were equal could not be rejected ($p < .05$) although preliminary experiments indicated this might be the case (Curry and Govindaraj, 1976). Threshold values for different nominal conditions were also compared and it was found that changes from the nominal (T=3.0,Z=.2) yielded consistently higher thresholds for both variance and bandwidth than the other three nominal processes (Table 1). The reasons for this difference can not be explained using simple velocity concepts since the RMS velocity for the processes considered here is given by

$$\sigma_{\dot{x}}^2 = \omega_n \sigma_x \qquad (4)$$

which is independent of the damping ratio.

Detection Times

Since the subjects who took part in these experiments went through the previous series of experiments for threshold determination, they had extensive experience with detection of parameter changes in a random process. In addition, the experiment was conducted over a period of several weeks and it is therefore reasonable to assume that the subjects were 'well-trained observers'.

The detection time data for changes in variance and bandwidth are plotted in Figures 1-4 in the form of log detection time vs, the log of the change in RMS velocity. (Also included for later reference are model predictions.) All curves indicate that using the change in RMS velocity is a reasonable method for summarizing the data. Although there is a tendency for increases in variance and bandwidth to be detected more quickly than decreases, the variability in responses over the population of observers suggests that these differences are of little practical

significance even if statistical significnace were to be achieved.

Models For Failure Detection

General Description

The process displayed to the subject consists of up and down motion of a horizontal line and the subject observed the position of the line continuously, but it is well known that a human is also sensitive to the velocity of the motion being observed. Initial analysis showed that if velocity were to form the basis of the detection task, more accurate predictions could be made of the experimentally observed detection times. Hence, a scalar observation, consisting of the rate of motion, was considered as the observed variable. The observations are assumed corrupted by an additive noise v(t), which was modeled as a zero mean Gaussian process. The input to the failure detection model thus consists of the observation (display rate) plus Gaussian white noise. The failure detector is modeled here as an estimator to generate a sequence of uncorrelated residuals which are utilized in the decision mechanism.

For all the models, the states were updated at 5/60 ths of a second, and the stimulus presentation was the same as in the experiment. The summations needed for the decision function were done with a first order filter with a long time constant, starting at 5 seconds after the start of the trial (1/(Ts+1), T = 0.001). Thus it is effectively a direct summation. In all cases, P(fa) = 0.05 was used for setting the decision boundaries.

Residual-Variance Detector

A Kalman filter designed for the nominal condition was used for the estimation portion of the model to detect changes in both frequency and variance. There are now two separate problems: (1) the detection of failures in frequency, and (2) the detection of failures in variance. The mean of the residual remains zero for both changes, since the system (the shaping filter), and the Kalman filter are linear, and the overall input is a zero mean white Gaussian process. However, the variance of the residual changes for failures of both types. This characteristic of the residual motivated a modeling approach using residual variance change as a failure indicator.

As with previous models of human failure detection, we use a sequential probability ratio test on the residuals with the further constraint that each time the decision function confirms normal operation, it is reset to zero (Chien and Adams,1976, Gai and Curry, 1976). For increases in residual variance, this leads

to a decision function of the form:

$$\lambda_m = \begin{cases} \bar{\lambda}_m, & \bar{\lambda}_m \geq 0 \\ 0, & \bar{\lambda}_m < 0 \end{cases} \qquad (5)$$

$$\bar{\lambda}_m = \lambda_{m-1} - \log\left[\frac{\sigma_1}{\sigma_0}\right] - \left[\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\right]\frac{r_m^2}{2} \qquad (6)$$

$$\text{if } \geq A_1, \qquad \text{decide } H_1$$

where r_m is the observed residual at time m, σ_0 is the nominal residual variance, σ_1 is the "failed" value of the residual variance ($\sigma_1 > \sigma_0$) and the constant A_1 is chosen to provide the specified probability of false alarm (Gai and Curry, 1976). Equations similar to (5,6) can be derived for an hypothesized residual variance decrease. Two such decision mechanisms operating in parallel (one for a residual variance increase, one for a decrease) were applied to the random processes experienced by the observers; the mean model detection time is shown in Figures 1-4. The parameters were chosen for each type of failure to represent the population's detection time rather than to represent each individual's performance.

The agreement between model and data is good, even to the point of mimicing differences resulting from nominal bandwidths differences.

Velocity Magnitude Discriminator

In this approach the velocity of the line is used to test for the means in the following manner. Since the velocity itself is a zero mean process, the magnitude of velocity, i.e., without regard to the direction in which it is moving, is taken as the basis for the model. In the initial learning phase, the Kalman filter is used to obtain an estimate of the mean speed. After this estimate is made , the Kalman filter stage is 'shut off', and the second stage is used as a comparator. This compares the observed speed (velocity magnitude) with the estimated mean value and generates the error residuals. Under normal conditions with no failure, this is a zero mean process. But when a failure does occur, the mean speed changes, and this is reflected in the mean of the residuals. A modified sequential probability ratio test as described above is then used to detect changes in mean speed. Detection time results of this model compared as favorably with the data as the variance detection model described above.

Comparisons Among Models

An important motivation for trying to determine models other than the one used for variance changes was the observation that the higher stimulus values predicted very low detection times. (Low detection times from the model seem reasonable if the human's reaction time is taken into account. This may be taken to be in the range 0.2-0.3 seconds.) Also, it was interesting to test if velocity could be used without regard to its sign. For the cases under consideration, both the models appear to perform well. A third approach was also tried, based on the idea that the subject might be estimating the average number of zero-crossings or level crossings to obtain an estimate of frequency. This model performed well for failures with an increase in frequency, but decreases had a very high false alarm rate compared with the subjects. A different decision criterion that accounts for the subject's prior information that the failure occurs at least after 8 seconds after starting might give fewer false alarms. A more detailed investigation is necessary to test the validity of this 'zero-crossing detector' model.

In view of the above comments the Kalman Filter/Sequential Decision model appears to be the more parsimonous choice because it accommodates both variance and bandwidth changes in one framework. Moreover, it is consistent with the model philosophy previously established for changes in means. A major difference yet to be resolved is the fact that display position input is used in the detection of changes of mean, and display velocity is used in the detection of changes of variance and bandwidth.

References

Chien, T.T., & Adams,M.B., A sequential failure detection technique and its application, IEEE Trans. Auto. Control, October 1976

Curry, R.E. & Gai, E., Detection of random process failures by human monitors, in Monitoring Behavior and Supervisory Control (Sheridan & Johannsen, Ed) New York: Plenum Press, 1976

Curry, R.E., & Govindaraj,T., The psychophysics of random processes, Proc. 12th Ann. Conf. Man. Control, 1976

Gai, E. & Curry, R.E., A model of the human observer in failure detection tasks, IEEE Trans. Syst., Man & Cybernetics, Feb 1976

Govindaraj,T., Failure detection by human observers, S.M. Thesis, Departemnt of Aeronautics and Astronautics, MIT, Cambridge, MA, 1977

Kleinman, D.L., & Curry, R.E., Some new control theoretic models for human operator display monitoring, IEEE Trans. Sys., Man Cybernetics, November 1977

Levison,W., A control theory model for human decision making, Proc. 7th Ann. Conf. Manual Control, 1971

Phatak,A.V., & Kleinman, D.L., Current status of models for the human operator as a controller and decision maker, Proc. AGARD Conf. No. 114, Oct 1972

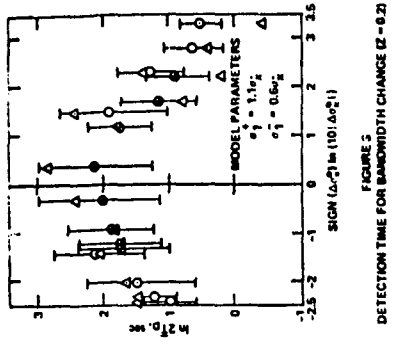


FIGURE 3
DETECTION TIME FOR BANDWIDTH CHANGE (Z = 0.2)

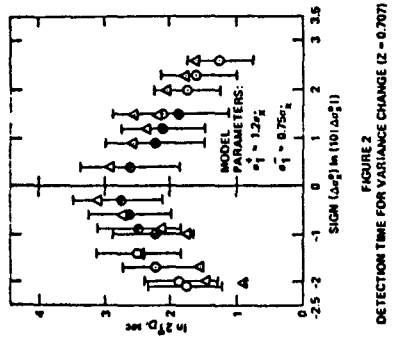


FIGURE 2
DETECTION TIME FOR VARIANCE CHANGE (Z = 0.707)

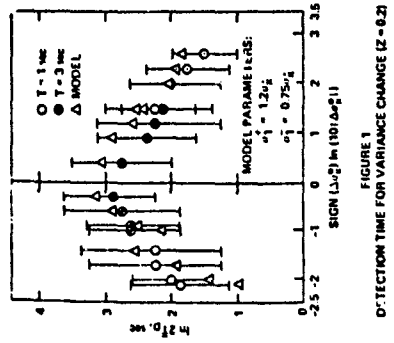


FIGURE 1
DETECTION TIME FOR VARIANCE CHANGE (Z = 0.2)

(1) Change in period

Nom	Th (log)	R	W	W-W$_0$	%Nm
12	-0.078 (0.051, 91	0.924	6.800	0.517	8.2
	0.064 (0.048, 91	1.066	5.696	-0.387	6.2
17	-0.076 (0.046, 91	0.927	6.778	0.495	7.9
	0.067 (0.050, 91	1.080	5.764	-0.519	8.3
22	-0.138 (0.090,12)	0.871	2.405	0.310	14.8
	0.124 (0.059,12)	1.131	1.852	-0.243	11.6
27	-0.085 (0.060,12)	0.917	2.284	0.188	8.9
	0.108 (0.053,12)	1.114	1.881	-0.214	10.2

(2) Change in variance

Nom	Th (log)	R	%Nm
12	-0.166 (0.087,10)	0.847	15.3
	0.177 (0.097,10)	1.190	19.4
17	-0.183 (0.101, 8)	0.825	17.5
	0.152 (0.101, 8)	1.162	16.2
22	-0.279 (0.090, 8)	0.757	24.3
	0.345 (0.168, 8)	1.410	41.0
27	-0.210 (0.078, 9)	0.811	18.9
	0.191 (0.085, 9)	1.211	21.1

Nom - Nominal timeout and damping ratio
Th (log) - Threshold on log units (in $(F/g)^2$)
R - Ratio of the parameter to nominal
W - Frequency (rad/sec)
W-W$_0$ - Difference from nominal
%Nm - Percent threshold (period or variance)
Standard deviation and number of subjects are shown in parentheses.

TABLE 1
THRESHOLDS FOR PERIOD BANDWIDTH AND VARIANCE CHANGES

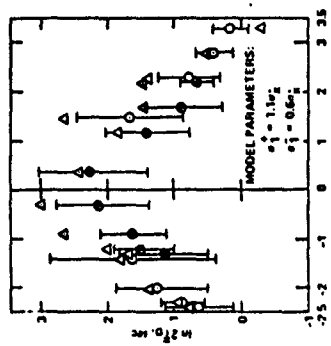


FIGURE 4
DETECTION TIME FOR BANDWIDTH CHANGE (Z = 0.707)

N79-17497

A QUEUEING MODEL OF PILOT DECISION MAKING IN A MULTI-TASK FLIGHT MANAGEMENT SITUATION*

R. S. Walden and W. B. Rouse

ABSTRACT

Allocation of decision making responsibility between pilot and computer is considered and a flight management task, designed for the study of pilot-computer interaction, is discussed. A queueing theory model of pilot decision making in this multi-task, control and monitoring situation is presented. An experimental investigation of pilot decision making and the resulting model parameters are discussed.

INTRODUCTION

Automation has found increasing use in aircraft, not only with respect to mechanical systems, but with respect to jobs normally thought of as belonging to the airplane pilot. As airplane numbers increase and navigation becomes more complex, the trend will certainly continue. Wempe [1] suggests that an advanced automated navigation system will become necessary in the future, because the required control performance will exceed human capability. However, he presents two reasons why completely automatic control of aircraft is unlikely in the near future: a machine required to replace a pilot would be quite costly; furthermore, machines do not adapt well to unanticipated events. Control of aircraft thus will be the task of a man-machine system.

With the recent rapid expansion of digital technology (especially microprocessors), the words "man-machine" and "man-computer" have become practically synonymous. The computer can have several roles in an airborne man-computer system. Special-purpose processors can be dedicated to such things as autopilot and navigation aids, displays, communication systems, or subsystem monitoring; in such applications, the computers would serve merely as digital controllers of the various devices. The pilot's workload would thus be reduced as the computers handle much of the low-level information processing.

Or, to go a step further, the computer could have a more general role. The computer might have inputs from a variety of aircraft devices and subsystems, and increased control capabilities as well. The computer would then be expected to initiate actions of its own (i.e., make decisions), based on the inputs it receives. The pilot's workload would be reduced at a higher level, by the computer sharing in decision-making tasks. It is this latter concept of the role of the computer in airborne systems which is the concern of the research described herein.

A computer with significant decision-making responsibilities in the flight-management task can create potential problems, however. One of the most crucial problems is that of allocation of decision-making responsibility between pilot and computer [2]. Allocation can either be static (situation-independent), or dynamic (situation-dependent).

Consider first a system in which allocation of responsibilities is static. There is a pre-defined, non-overlapping set of responsibilities for each decision-maker (man and computer). Such a system has one desirable feature in that no confusion is possible over who should perform a given task. However, there are several serious drawbacks to the static allocation of responsibility. These include:

1. Poor utilization of system resources due to each decision maker being prohibited from performing the other decision maker's tasks, even when one decision maker is overloaded and one is idle.

2. Difficulty for the pilot when unforeseen circumstances force him to take over tasks for which he normally has no responsibility.

3. Possible pilot resentment at not being allowed to perform certain tasks, especially when he has nothing else to do.

For the above reasons, a dynamic allocation of responsibility might be more appropriate. Simulation studies of this alternative by Rouse [3] support this conclusion. Succinctly, the idea is that the server who is most able, in a given situation, will assume responsibility for a task.

Dynamic allocation of responsibility is not without its drawbacks, however. The most important problem to be solved is that of communication between pilot and computer. This is necessary to avoid conflicts which occur when the two decision makers act independently, resulting in degraded performance. (Rouse [3] presents simulation results that illustrate this effect). A possible solution to this problem is to use the computer as a backup for the pilot, to be turned on when the pilot's workload becomes too great and/or his performance degrades.

* This research was supported by the National Aeronautics and Space Administration under NASA - Ames grant NSG-2119.

The man-computer system is thus seen as one in which both the pilot and computer monitor each other--the pilot supervises and monitors the computer's activities, for the purpose of insuring that the computer is performing adequately, and that no malfunction has occurred; the computer monitors the pilot to determine when he is overloaded and needs assistance. Such a system would hopefully minimize conflict and avoid the difficulties discussed above.

In order for the computer to be able to tell when the pilot needs assistance, it must first have a method for determining the pilot's workload, in terms of task demands. Then, it needs a criterion for deciding if it should request responsibility for any tasks. Chu and Rouse [4] have proposed a formulation of this problem. To apply this formulation to any specific situation, the first step is development of a model of the unaided performance of the human with respect to tasks typical of those encountered in the situation of interest.

The purpose of this report is to present the results of pursuing this first step in the context of flight management. An experimental situation which was developed to study pilot performance in simulated flight conditions is described. Alternative models that describe pilot decision making in this flight management situation are presented. Finally, an experiment is discussed whose purpose was investigation of the suitability of the alternative models.

THE TASK

The pilot is presented with a simulated airplane instrument panel, drawn on a CRT (see Figure 1). The display includes standard aircraft instruments, such as artificial horizon, altimeter, heading and airspeed indicators. Also displayed is a map which indicates the course the airplane is to follow. An airplane-shaped symbol, superimposed on the map, indicates the airplane's actual position. A small circle moves along the mapped course at a constant speed of 600 fps (183 m/s) and indicates the position the aircraft should have for it to be on course and schedule. Near the lower edge of the display, several dials are shown, which represent (abstractly) gauges for such things as fuel, electrical, or hydraulic subsystems. The gauge pointers move with apparently random motion, generated by passing Gaussian white noise through a second-order sampled-data filter.

The pilot controls Boeing 707 aircraft dynamics as discussed by Blakelock [5]. He controls the aircraft pitch and roll dynamics with a joystick. Another control stick controls the airspeed. The pilot's task is to fly the airplane along the mapped route, maintaining a fixed altitude and stable pitch and roll attitude. The pilot is supposed to keep as close to the "on-schedule" dot, and as close to the mapped course, as possible, while holding the aircraft at 40,000 ft. (12,192 m.) altitude.

In addition to flying the airplane, the pilot monitors the subsystem indicators for possible events. (An event has occurred when the pointer motion slows and stops, pointing downward as shown for subsystem 4 in Figure 1.) When the pilot thinks an event has occurred, he enters the corresponding number on the keyboard. (If more than one event has occurred, the pilot chooses the higher-priority, or lower subsystem number, event.) The display shown in Figure 2 then appears.

On this new display, the subsystem dials have been replaced with two rows of numbers, one labelled 'BRANCH', the other labelled 'STATE'. This corresponds to the first level of a checklist associated with the subsystem which was selected. The pilot looks for a branch with a state of '0' and enters the branch number on the keyboard. If the checklist for that branch continues, the next checklist level is displayed. When the end of the checklist is reached, the subsystem indicators reappear, with the subsystem indicator motion restored. These actions simulate the checklist procedures a pilot might go through when performing tasks in a real aircraft.

The apparatus used for the experiments, the simulation software used to produce the above task, a description of the aircraft dynamics, and an explanation of the subsystems portion of the simulation are discussed in more detail in Walden's thesis [6].

THE MODEL

The role of the human operator in complex man-machine systems is changing from that of in-loop controller to that of supervisor or monitor of automatic, possibly computer operated, controls. In such a role, it is the human's monitoring performance which is of major interest. The assumption is that essential control tasks will be performed adequately, perhaps at the expense of monitoring performance.

Thus, the goal of the modeling effort reported here is to predict the subsystem task performance, measured in terms of mean and variance of subsystem event waiting time, as a function of task parameters 1) event arrival rate, and 2) control task difficulty.

Existing models prove inadequate in describing the multi-task flight management situation discussed here. Control models, for example the Kleinman, Baron, and Levison

optimal control model [7] (and its many extensions), are concerned only with predicting control performance; a model useful in predicting subsystem service performance is sought. Instrument-monitoring models do not adequately describe the multi-task flight management situation, either. The instrument-monitoring model of Senders [8] focuses on details of instrument scanning, such as instrument fixation frequencies and sequences. That of Carbonell [9] is concerned with sampling policy based on instrument priorities. Neither type of model is concerned with subsequent actions resulting from instrument observations.

A queueing theory model of the multi-task situation is proposed. The monitoring task is modeled as a single-server queueing system, with subsystem events as customers. The monitoring task is, in fact, a queueing system, and it seems natural to use a queueing theory description for it. The proposed model is a simulation model; an appropriate analytical solution is not available. (See Chu [10] for explanation of the basic operation of the queueing simulation.)

In the first modeling attempt (hereafter referred to as Model 1), the monitoring task is described as a $M/E_k/1$:HPRP/N/N queueing system. (See White, Schmidt, and Bennett [11] for definitions of the standard queueing terminology employed here.) The control task is not explicitly modeled. Instead, the control task is included implicitly in the service time distribution parameters (i.e., for the Erlang distribution, the mean and shape parameter k), which change with increased map complexity (due to the control task). Computer simulation of this queueing system provided preliminary data useful in making later improvements, but did not match the experimental data well under low traffic conditions. (See RESULTS)

The second refinement of the model (referred to as Model 2) attempts to account for the impact of the control task on the subsystem task performance. This is done by using a higher effective arrival rate for subsystem events:

$$\lambda_{effective} = \lambda_{actual} + \Delta\lambda$$

where $\Delta\lambda$ is the increase in arrival rate needed to match empirical subsystem event waiting time. The control task is thus incorporated into the monitoring portion of the task as increased event arrivals in the subsystem queues. Using the same simulation program as before, subsystem event arrival rate was varied until the mean waiting times predicted by the simulation matched the experimentally observed waiting times. This model provides a good prediction of the waiting time distributions for individual subsystems (see RESULTS). However, in terms of the physical system, the $\Delta\lambda$ parameter is admittedly unappealing.

With the third refinement, the model (Model 3) approaches the real situation most closely. The control task is incorporated as a special queue; customers arriving in the control task queue can preempt the service of a subsystem event. Service of the subsystem event will then be resumed at the point of interruption after the control action is complete. This queueing system is described as a $M/E_k/1$:PRP/N/N queue.

One of the problems with this third model is defining a "customer" for the control task queue. In fact, it is displayed error (more specifically, increments of displayed error) which queues (accumulates) for attention; however, the size of the incremental "customer" is unknown. A reasonable approach is to consider the assumption made earlier that the control task customer preempts subsystem service; i.e., whenever a control task customer is present, subsystem monitoring (or event service) is immediately preempted, and the control task served. A customer in the control task queue could then be defined (loosely) as a "significant" or "action-evoking" amount of display error; when a significant error is present, subsystem service will be preempted and control inserted to null the error. A simple way to indirectly measure the frequency with which "significant" errors arrive (queue) is to measure the frequency with which displayed errors are serviced, i.e., the frequency at which control actions are inserted to null them.

To estimate the frequency of control responses, the number of separate control actions performed by the pilot can be counted, and from this number, an arrival rate for control task customers can be obtained. To obtain an estimate of the average service rate for a customer in the control task queue, the control task service rate parameter in the simulation program can be varied until a good fit to experimental data is obtained.

To account for the increased waiting time incurred from errors such as false alarms and incorrect actions, modifications to the basic simulation model were made. False alarms are accounted for by using another queue, with event (false alarms) arrival rate equal to the arrival rate of false alarms in the experiment. For the task discussed here, the service time used for a false alarm is one-fifth the mean event service time, since only the first of five levels of the checklist is shown when a false alarm occurs.

Incorrect actions are accounted for in the simulation by first estimating the probability of an incorrect action while servicing a subsystem event. This is done simply by dividing the number of incorrect actions by the number of events. Based on this rate, incorrect actions are simulated. For the monitoring task discussed here, the penalty for an incorrect action is an increase in waiting

time equal to one-half the mean service time for subsystem events. (On the average, if incorrect actions occur randomly, the pilot will be halfway through a checklist procedure, and thus waste one-half of the average service time.) The waiting-time penalty could be even greater, if after an incorrect action (and the subsystem indicators are returned to the display), the pilot detects and services an event in a higher-priority queue. This is also accounted for in the simulation model.

AN EXPERIMENT

Using the experimental situation described earlier, an experiment was performed to study subject "pilot" performance in a multi-task flight management situation. The two independent variables in the experiment were the inter-arrival times of subsystem events, and the difficulty of the flight path to be followed by the pilot.

In the first part of the experiment, only the subsystem monitoring task was considered. An "autopilot" kept the airplane on course, coincident with the "on-course indicator" marker. 10-minute trials, using 30, 60, and 90 second average interarrival times (per subsystem), were run with each of six subjects. (The actual interarrival times were exponentially distributed about the mean.) To equally distribute the effects of practice obtained during a sequence of trials, a balanced design was employed, such that the three trials were given in a different order to each subject (see Table 1). Six subjects participated, all of which were male students or former students in engineering.

During this first part of the experiment, two sessions, three ten-minute trials in each, were run with each subject. The first session served to train the subjects for the monitoring task; little improvement in performance was noted during the second session. (Only the data from the second session was used for analysis.)

Data from this part of the experiment consisted of the following:

1. Time of occurrence of a subsystem event
2. Time of response to the event
3. Time of completion of diagnostic action for the event.

Failures to complete diagnostic action after an initial response (incorrect actions) were also noted, as well as responses to nonexistent events (false alarms).

The second part of the experiment consisted of both monitoring and control tasks. After considerable training, the subjects participated in four formal experimental sessions, each of which involved two trials of about 15 minutes duration. For the first trial of a session, the subject was given a simple map (few turns) to follow (Figure 3). In the second trial, the course was more complex (Figure 4). The mean event interarrival time for the monitoring task remained the same during each session, but varied from session to session, as indicated in Table 1.

To establish baseline performance for the control task, the first session included only the aircraft control, but no monitoring. The remaining sessions were run with the three levels of monitoring workload used earlier. The balanced design for varying the monitoring task workload was also used in these sessions (Table 1). Prior to performing the final trials, the subjects practiced "flying" the airplane simulator, both without and with monitoring tasks. When they were able to fly the complex map, with the lowest (30 sec.) subsystem event interarrival time, and also maintain control of aircraft attitude and position, the final trials were run.

Data obtained from this part of the experiment (sampled every 0.25 sec.) included the three measures listed above, as well as control task information:

4. Aircraft position and attitude
5. Pilot control inputs.

Performance data for four levels of monitoring workload and three levels of control task workload were thus obtained from the experiment as a whole.

RESULTS

The raw data obtained during the experiment was analyzed to obtain subsystem event service time and waiting time statistics, summarized in Table 2. Errors (false alarms and incorrect actions) were also counted, and are shown in Tables 3 and 4. Table 5 shows three measures of control task performance (RMS perpendicular distance from the course, RMS pitch and roll angles), for the subsystem event arrival rates and maps used in the experiment (also presented graphically in Figure 5.)

Subsystem event service time is measured from initiation of diagnostic action to the completion of the action, provided no other subsystem action intervenes (after an incorrect action, a different, higher-priority subsystem event could be serviced). Figure 6 shows the empirical service time statistics. The average service time appears

to be independent of subsystem event arrival rate. As discussed in THE MODEL section, the preemption of subsystem service by the control task is reflected as an increase in the subsystem service time. This increase is apparently not a function of control task difficulty, but simply of whether or not the control task is present.

The subsystem event waiting time is measured from the time of occurrence of an event to the time at which diagnostic action for that event is completed. Waiting time is the sum of two terms: response time, the time from event occurrence to initiation of diagnostic action; and service time, from initiation to completion of diagnostic action. As shown in Figure 7, waiting time increases with subsystem event arrival rate. This is due to the queueing effect on the response time.

As expected, waiting time also increases when the control task is added. This is due in part to the increased service time when the control task is present. However, larger service time alone does not account for all of the increase in waiting time. The control task also affects the response time. This is the queueing effect of the control task "customers" on the subsystem service performance. This effect is also only a function of the mere presence of the control task, rather than the control task difficulty, as was the effect on service time.

False alarms (Table 3) tend to occur less frequently with increasing control task difficulty and event arrival rate. When the control task was present, few false alarms were made; many of those which occurred were probably simply typing mistakes made in response to real subsystem events (this could not be detected). False alarms wasted on the order of 1 sec. of time, and probably had insignificant effect on subsystem task performance.

Incorrect actions (Table 4) tended to increase with subsystem event arrival rate, but showed no consistent variation with control task difficulty. Incorrect actions are significant because, on the average, one-half the average service time is wasted. Furthermore, once an incorrect action is made, the pilot can switch to a higher-priority event, without completing the diagnostic already begun. This contributes unevenly to the waiting time across the subsystems; such an occurrence during service of subsystem 6, for example, might result in that diagnostic action being delayed considerably, while similar preemption during service of subsystem 2 would not result in as long a delay. The impact of the incorrect actions is thus to increase both the average and standard deviation of waiting time for lower-priority processes.

Each data item from the results of the simulation to be discussed here is based on 10,000 simulated events. As might be expected, for all three sets of simulation results, the computer simulation is more consistent (smaller standard deviation) than the experimental subjects.

Table 6 shows both the empirical subsystem waiting time and the results from simulation Model 1. Input parameters to simulation Model 1 were:

1. Probability of incorrect action, given that a subsystem event has occurred. This was approximated from the empirical data by dividing the number of incorrect actions by the number of events.

2. False alarm arrival rate and service rate. The service rate was assumed to be five times the average subsystem event service rate for the task, because only one of the five levels of the checklist need be scanned to realize a false alarm has been made. The arrival rate for false alarms was obtained by dividing the number of false alarms by the total time during which false alarms could occur. (The total time during which false alarms could occur was equal to the total elapsed time minus the total time spent servicing false alarms, since false alarms could not occur while one was being serviced.)

3. Subsystem event arrival rate, one of the independent variables in the experiment.

4. Average subsystem event service rate, obtained from the empirical service time data.

5. Erlang service time distribution shape parameter k, the square of the ratio of mean to standard deviation. This was calculated from the empirical service time data (Table 2). This value averaged 19 for the monitoring-only tasks, and was 3 for all other tasks.

A subsystem event could not be detected immediately. By averaging the minimum times from event occurrence to initial response which were recorded in the experimental trials, 4.5 sec. was estimated as the average time after occurrence of an event before it could be detected. In the computer simulation, events are detected immediately; to account for the detection time delay, 4.5 sec. is added in the simulation to the basic waiting time, to give the adjusted waiting time.

Summarizing the results from Model 1, when no control task is present, the model predictions match the empirical data well. Both mean and standard deviation of waiting times for individual subsystems are close. However, when the control task is present, a close match is obtained only

for the high subsystem arrival rate condition. With lower arrival rates, the model predictions are smaller than the empirical results. This can be explained in the following way. With high arrival rate, events occur frequently and most control actions preempt subsystem diagnostics in progress. The control task effect is therefore almost entirely included in the increased service time for subsystem events. With lower arrival rates, a significant portion of the control actions can be performed without preempting subsystem diagnostics, but may increase the time to respond to a subsystem event. Thus, for the lower subsystem event arrival rates, the control task is not adequately modeled as simply an increase in service time for subsystem events.

Simulation Model 2, which uses an effective arrival rate (see THE MODEL) to account for the control task, produced a good match to empirical data (see Table 7). The simulation input parameters are the same as those above, except that an effective arrival rate of subsystem events $\lambda_{effective} = \lambda_{actual} + \Delta\lambda$ is used instead of the actual arrival rate. $\lambda_{effective}$ was obtained by varying $\Delta\lambda$ until a close fit to the empirical waiting time data was found. (Since Model 1 produced an adequate fit to empirical data for the high arrival rate conditions, the effective arrival rate was only computed for low and medium arrival rates, both simple and complex maps.) The various values of $\Delta\lambda$ were averaged (to give 0.0095), and this average used to calculate the effective arrival rates used to produce the results summarized in Table 7.

The results of simulation Model 3 are shown in Table 8. The probability of incorrect action, false alarm arrival rate and service rate, and subsystem event arrival rate are the same parameters as for Model 1. The subsystem event service rate was the average of the three rates obtained from the baseline monitoring-only tasks. The Erlang shape parameter k was the average of the three values obtained for the monitoring-only tasks.

The arrival rate of customers in the control task queue λ_c was calculated from the number of distinct control actions counted (N_c), the elapsed time of the corresponding trials (T_e), and the average duration of the control actions (T_d) (all empirical quantities) using:

$$\lambda_c = \frac{N_c}{T_e - N_c * T_d}$$

(Actually, only aileron control inputs were counted; elevator inputs were infrequent and of short duration compared to aileron inputs. See Figure 8 for an example of aileron control input vs. time for one trial.) The service rate for the control task queue was varied until close fits to empirical waiting time averages (across all subsystems) were obtained. The service rates for simple map were then averaged, as well as those for the complex map. An Erlang shape parameter k of 2 was used for the control task queue service time distribution. These values were used to produce the results shown in Table 8.

It is important to note that the results obtained for this third model were based on the assumption that control task service rate does not vary with subsystem event arrival rate. This assumption may be inappropriate. However, if control task service rate varies with both control task difficulty and subsystem event arrival rate, one could arbitrarily match almost any data. What is needed is a method of determining the control task service rate directly from the control signals. This is the most important extension needed by this queueing model.

SUMMARY AND CONCLUSIONS

The suitability of three alternative queueing models of the flight management situation was investigated, and the results presented. Attention was focused on predicting pilot performance in the subsystem monitoring task. In the first model, the subsystem monitoring task was explicitly modeled as a $M/E_k/1$:NPRP/N/N queueing system. The control portion of the task was included in the empirical subsystem event service time, which increased when the control task was present. This model adequately predicted the subsystem event waiting time performance only for the tasks with highest arrival rates.

The second model accounted for the control portion of the task (with low and medium arrival rates) by using an effective arrival rate for subsystem events. By adding an increment of arrival rate to the actual event arrival rate, the empirical waiting time performance could be predicted satisfactorily by computer simulation. However, using an effective arrival rate leaves much to be desired from the point of view of what is physically happening in the overall task.

The third model incorporated the control task as a separate, special queue which could preempt subsystem event service whenever a customer arrived. The task was thus modeled as a $M/E_k/1$:PRP/N/N queueing system. An estimate of the arrival rate for control task queue customers was made. The service rate for the control task was used as a fitting parameter. By varying this parameter in the simulation, a reasonable fit to empirical data was obtained.

The three models presented represent successive refinements in an attempt to incorporate the control task into the queueing model of the flight management situation. The subsystem monitoring task was easily formulated using a queueing description, and its basic representation was the same in all three model refinements. The third model

refinement most closely approaches the real situation with respect to representation of the control and subsystem monitoring tasks; it is therefore intuitively appealing (although the control task parameters are not well-defined).

The goal of this research was to determine if a queueing model would provide an adequate description of pilot decision making in a multi-task, control and monitoring situation. Reasonable predictions of performance of the subsystem task were obtained with little or no fine tuning of parameters. This suggests that the model is promising, and can be successfully employed not only in the adaptive computer aiding scheme mentioned earlier, but perhaps in a broader class of similar multi-task situations.

For example, the proposed queueing model could be fit to multi-task situations and then the fractions of attention required by each task could easily be determined since they are inherent outputs of a queueing formulation. The advantage of this approach is that the fractions of attention are no longer free parameters as they are in control theory models of human decision making in multi-task situations [12, 13].

REFERENCES

1. T. F. Wempe, "Flight Management - Pilot Procedures and System Interfaces for the 1980-1990's," AIAA paper No. 74-1297, AIAA Life Sciences and Systems Conference, November 1974.

2. W. B. Rouse, "Design of Man-Computer Interfaces for On-Line Interactive Systems," Proceedings of the IEEE, Special Issue on Interactive Computer Systems, Vol. 63, No. 6, pp. 847-857, June 1975

3. W. B. Rouse, "Human-Computer Interaction in Multi-Task Situations," IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-7, No. 5, pp. 384-392, May 1977.

4. Y. Chu and W. B. Rouse, "Optimal Adaptive Allocation of Decision Making Responsibility Between Human and Computer in Multi-Task Situations," to appear in the Proceedings of the 1977 International Conference on Cybernetics and Society, September 1977.

5. J. H. Blakelock, Automatic Control of Aircraft and Missiles, New York: John Wiley & Sons, Inc., 1965.

6. R. S. Walden, "A Queueing Model of Pilot Decision Making in a Multi-Task Flight Management Situation," MSME thesis, University of Illinois at Urbana-Champaign, July 1977.

7. D. L. Kleinman, S. Baron, and W. H. Levison, "A Control Theoretic Approach to Manned-Vehicle Systems Analysis," IEEE Transactions on Automatic Control, Vol. AC-16, No. 6, pp. 824-832, December 1971.

8. J. W. Senders, "The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems," IEEE Transactions on Human Factors in Electronics, Vol. HFE-5, No. 1, pp. 2-5, September 1964.

9. J. R. Carbonell, "A Queueing Model of Many-Instrument Visual Sampling," IEEE Transactions on Human Factors in Electronics, Vol. HFE-7, No. 4, pp. 157-164, December 1966.

10. Y. Chu, "Optimal Adaptive Allocation of Decision Making Responsibility Between Human and Computer in Multi-Task Situations," Ph.D. thesis in progress, University of Illinois at Urbana-Champaign.

11. J. A. White, J. W. Schmidt, and G. K. Bennett, Analysis of Queueing Systems, New York: Academic Press, Inc., 1975.

12. R. E. Curry, D. L. Kleinman, and W. C. Hoffman, "A Design Procedure for Control/Display Systems," Human Factors, Vol. 19, No. 5, October 1977.

13. S. Baron and W. H. Levison, "Display Analysis with the Optimal Control Model of the Human Operator," Human Factors, Vol. 19, No. 5, October 1977.

Table 1. Experimental design

	Session Number									
	1	2		3		4		5	
	No control task	Control Task		Control Task		Control Task		Control Task	
Subject		Simple Map	Complex Map	Simple Map	Complex Map	Simple Map	Complex Map	Simple Map	Complex Map
1	30* 60 90			30	30	60	60	90	90
2	30 90 60			30	30	90	90	60	60
3	60 30 90	No Monitoring		60	60	30	30	90	90
4	60 90 30			60	60	90	90	30	30
5	90 30 60			90	90	30	30	60	60
6	90 60 30			90	90	60	60	30	30

* Table entries are mean subsystem event interarrival times (sec.)

Table 2. Subsystem event service performance (empirical)

Subsystem No.	Subsystem event arrival rate (events/sec) 0.01111					0.01667					0.03333				
	N	T_s	σ_s	T_w	σ_w	N	T_s	σ_s	T_w	σ_w	N	T_s	σ_s	T_w	σ_w
						No Map									
1	18	5.55	0.81	12.97	3.06	82	5.57	1.32	12.86	3.67	84	5.96	2.08	14.67	4.20
2	25	5.84	1.62	12.10	3.13	88	5.40	1.10	14.11	4.28	82	5.42	1.19	14.68	3.93
3	30	5.36	0.58	12.38	2.51	86	5.28	1.21	14.43	5.36	82	5.29	1.31	17.87	7.25
4	24	6.22	2.86	12.80	4.67	36	5.27	0.67	15.91	5.88	90	5.44	1.35	18.69	10.07
5	32	5.41	1.39	12.62	5.74	42	5.39	0.62	15.74	6.38	58	5.78	1.78	24.26	18.38
6	48	5.58	0.91	13.55	3.82	35	5.37	0.90	17.00	7.56	52	5.39	1.25	32.88	27.59
All*	177	5.63	1.38	12.82	4.01	289	5.41	1.07	14.62	5.48	448	5.55	1.54	19.42	14.10
						Simple Map									
1	34	8.59	6.53	19.17	6.79	44	7.92	3.87	19.69	7.26	56	8.56	4.59	20.85	6.95
2	43	8.88	6.23	18.98	8.69	53	8.49	5.22	20.56	9.75	79	9.60	7.04	23.42	9.58
3	18	7.51	4.95	23.93	8.46	54	10.01	7.59	24.91	12.35	79	9.74	5.94	27.81	12.34
4	54	9.57	5.80	24.87	12.50	51	8.74	5.20	28.21	18.15	67	8.45	4.69	32.68	21.43
5	39	9.54	6.87	31.43	28.63	52	8.33	4.23	35.97	37.04	42	9.82	5.83	54.08	40.45
6	31	10.27	6.95	34.18	26.39	34	8.55	4.43	55.67	55.40	30	7.03	2.13	109.39	151.78
All	219	9.21	6.28	25.24	17.28	288	8.71	5.33	29.52	28.81	353	9.06	5.59	36.71	53.13
						Complex Map									
1	40	10.87	7.47	22.21	8.57	41	9.37	5.33	22.78	10.24	54	8.45	4.94	21.75	10.39
2	41	9.57	5.29	20.74	8.19	49	10.93	6.94	22.46	9.47	86	9.32	6.88	22.58	9.37
3	35	8.39	4.36	25.10	14.80	52	9.28	6.56	27.03	17.10	74	8.97	5.61	26.89	14.33
4	46	8.70	5.49	24.10	12.63	53	7.90	3.84	29.94	21.15	71	8.81	5.17	36.82	29.11
5	52	9.48	7.61	32.54	27.72	47	9.94	6.02	39.06	40.22	56	8.80	5.08	50.19	56.27
6	18	9.52	6.15	47.80	34.61	35	9.71	5.50	52.76	91.51	30	8.09	3.29	88.50	66.71
All	232	9.42	6.25	27.06	19.97	277	9.49	5.79	31.44	39.50	381	8.88	5.50	35.18	37.00

* Cumulative statistic, across all six subsystems

N - Number of subsystems events recorded

T_s, σ_s - Event service time average, standard deviation (sec.)

T_w, σ_w - Event waiting time average, standard deviation (sec.)

Table 5. Control task performance (empirical)

Subsystem event arrival rate (events/second)						
	0		0.01111	0.01667		0.03333
Statistic						
No map						
Mean position error (ft.)	1489		1852	2218		2698
RMS position error (ft.)	1961		2391	3847		6135
Mean course error (ft.)	790		1074	1096		1186
RMS course error (ft.)	1268		1609	1732		1730
RMS altitude error (ft.)	2700		4616	4395		669%
RMS pitch angle (radians)	0.096		0.118	0.134		0.119
RMS roll angle (radians)	0.294		0.304	0.329		0.315
Simple Map						
Mean position error (ft.)	3958		3247	3443		5203
RMS position error (ft.)	5526		4081	4309		6731
Mean course error (ft.)	1626		1638	1667		1921
RMS course error (ft.)	2334		2491	2033		2611
RMS altitude error (ft.)	669%		3842	3402		5355
RMS pitch angle (radians)	0.158		0.147	0.136		0.171
RMS roll angle (radians)	0.427		0.428	0.415		0.406
Complex Map						

Table 3. False alarms

Subsystem event arrival rate (events/sec.)			
	0.01111	0.01667	0.03333
No map	51*	20	26
Simple map	9	6	6
Complex map	4	6	13

* Table entries are total number of false alarms of six subjects for the specified task.

Table 4. Incorrect actions

Subsystem event arrival rate (events/sec.)			
	0.01111	0.01667	0.03333
No map	17*	24	50
Simple map	26	19	39
Complex map	15	21	38

* Table entries are total number of incorrect actions of six subjects for the specified task.

Table 6. Comparison of empirical data and queueing Model 1 simulation results

Subsystem No.	Subsystem event arrival rate (events/sec.) 0.01111				0.01667				0.03333			
	T_{wm}	σ_{wm}	T_{we}	σ_{we}	T_{wm}	σ_{wm}	T_{we}	σ_{we}	T_{wm}	σ_{wm}	T_{we}	σ_{we}
					No Map							
1	11.86	2.51	12.97	3.06	11.88	2.33	12.86	3.67	13.28	2.97	14.67	4.20
2	12.16	2.92	12.10	3.13	12.19	2.68	14.11	4.28	14.14	3.82	14.68	3.93
3	12.42	3.39	12.38	2.51	12.50	3.21	14.43	5.36	15.69	5.72	17.87	7.25
4	12.72	3.65	12.80	3.67	13.30	4.27	15.91	5.88	18.27	8.86	18.69	10.07
5	13.32	4.63	12.62	5.74	14.02	5.87	15.74	6.38	22.22	14.03	24.26	18.38
6	13.79	5.38	13.55	3.82	14.91	6.86	17.00	7.56	29.54	25.83	32.88	27.59
All*	12.71	3.92	12.82	4.01	13.10	4.58	14.62	5.48	18.21	13.03	19.42	14.10
					Simple Map							
1	17.91	7.23	19.17	6.79	17.58	7.15	19.69	7.26	20.47	7.56	20.85	6.95
2	18.47	8.41	18.98	8.69	18.47	7.79	20.56	9.75	22.87	9.60	23.42	9.58
3	19.11	9.17	23.93	8.46	19.41	9.64	24.91	12.35	27.11	14.25	27.81	12.34
4	20.39	11.74	24.87	12.50	21.95	12.08	28.21	18.15	36.38	24.30	32.68	21.43
5	21.67	13.99	31.43	24.63	24.25	16.31	35.97	37.04	54.51	46.07	54.08	40.45
6	24.46	18.37	34.18	26.39	28.90	23.29	55.67	55.40	102.03	109.94	109.39	151.78
All	20.29	12.22	25.28	17.28	21.60	14.16	29.52	28.81	35.81	43.68	36.71	53.13
					Complex Map							
1	17.49	7.50	22.21	8.57	19.14	7.71	22.78	10.24	19.91	7.46	21.75	10.39
2	18.46	8.07	20.74	8.19	20.33	8.74	22.46	9.47	22.39	9.69	22.58	9.37
3	19.28	9.65	25.10	14.80	22.27	11.29	27.03	17.10	26.54	13.39	26.89	14.33
4	19.94	10.84	24.10	12.63	25.02	14.79	29.94	21.15	35.03	22.92	36.82	29.11
5	21.47	13.16	32.54	27.72	29.34	21.50	39.06	40.22	51.69	43.92	50.19	56.27
6	23.23	16.61	47.80	34.61	35.64	31.04	52.76	91.51	94.50	97.10	88.50	66.71
All	19.92	11.46	27.06	19.97	24.90	18.07	31.44	39.59	34.54	40.26	35.18	37.00

* Cumulative statistic, across all six subsystems

T_{wm}, σ_{wm} - Model 1 event waiting time average, standard deviation (sec.)

T_{we}, σ_{we} - Empirical event waiting time average, standard deviation (sec.)

Table 7. Comparison of empirical data and queueing Model 2 simulation results

Subsystem No.	Subsystem event effective arrival rate (events/sec.) ($\lambda_{effective} = \lambda_{actual} + 0.0095$) 0.0206				0.0262			
	T_{wm}	σ_{wm}	T_{we}	σ_{we}	T_{wm}	σ_{wm}	T_{we}	σ_{we}
				Simple Map				
1	19.63	7.89	19.17	6.79	18.89	7.12	19.69	7.26
2	21.09	9.26	18.98	8.69	21.02	8.72	20.56	9.75
3	23.12	11.97	23.93	8.46	23.40	11.76	24.91	12.35
4	27.39	17.02	24.87	12.50	28.17	17.78	28.21	18.15
5	34.90	26.56	31.43	24.63	36.97	30.10	35.97	37.04
6	46.07	43.47	34.18	26.39	53.04	49.79	55.67	55.40
All*	27.59	23.01	25.24	17.28	29.39	25.45	29.52	28.81
				Complex Map				
1	19.80	7.90	22.21	8.57	20.26	7.60	22.78	10.24
2	21.83	9.61	20.74	8.19	22.77	10.01	22.46	9.47
3	24.19	12.29	25.10	14.80	25.96	12.97	27.03	17.10
4	27.83	17.37	24.10	12.63	32.92	21.51	29.94	21.15
5	34.42	24.97	32.54	27.72	43.34	36.55	39.06	40.22
6	45.38	42.74	47.80	34.61	67.87	69.66	52.76	91.51
All	27.95	22.71	27.06	19.97	32.33	32.58	31.44	39.59

* Cumulative statistic, across all six subsystems

T_{wm}, σ_{wm} - Model 2 event waiting time average, standard deviation (sec.)

T_{we}, σ_{we} - Empirical event waiting time average, standard deviation (sec.)

$\lambda_{effective}$ - Effective subsystem event arrival rate used in simulation

λ_{actual} - Actual subsystem event arrival rate used in experiment

Table 8. Comparison of empirical data and queueing Model 3 simulation results

Subsystem event arrival rate (events/sec.) (μ_s = 0.181, k_s = 19 for all tasks.)

Subsystem No.	0.01111				0.01667				0.03333			
	Simple Map (λ_c = 0.15, μ_c = 0.20, k_c = 2)											
	T_{wm}	σ_{wm}	T_{we}	σ_{we}	T_{wm}	σ_{wm}	T_{we}	σ_{we}	T_{wm}	σ_{wm}	T_{we}	σ_{we}
1	19.59	8.54	19.17	6.79	20.35	8.75	19.68	7.26	21.74	8.84	20.85	6.95
2	20.81	9.65	18.98	8.59	22.19	10.35	20.56	9.75	24.90	11.40	23.42	9.58
3	22.21	11.32	23.93	8.46	23.99	12.58	24.91	12.35	30.58	16.20	27.81	12.34
4	24.43	13.99	24.87	12.50	27.57	17.06	28.21	18.15	43.05	28.99	32.68	21.43
5	26.52	17.04	31.83	24.63	32.34	23.52	35.97	37.04	70.40	60.58	54.08	40.45
6	29.76	22.65	34.12	26.39	40.75	35.58	55.67	55.40	138.07	139.16	109.39	151.78
All[a]	23.74	14.86	25.25	17.28	27.33	20.59	29.52	28.81	41.21	53.78	36.71	53.13
	Complex Map (λ_c = 0.14, μ_c = 0.18, k_c = 2)											
1	20.27	9.21	22.21	8.57	21.20	9.45	22.75	10.24	22.05	9.48	21.75	10.39
2	20.88	10.02	20.74	8.19	22.82	10.82	22.86	9.47	25.21	11.88	22.58	9.37
3	22.37	11.75	25.10	14.80	25.05	13.39	27.69	17.10	30.75	16.88	26.89	14.33
4	23.21	13.26	24.10	12.63	29.76	19.09	29.94	21.15	42.80	31.38	36.82	29.11
5	26.14	16.88	32.54	27.72	34.38	26.04	39.06	40.22	70.36	60.26	50.19	56.27
6	29.33	21.63	47.80	34.61	43.33	37.52	52.75	91.51	141.03	152.17	88.50	66.71
All	23.64	14.69	27.06	19.97	28.85	22.26	31.44	39.59	41.53	56.26	35.18	37.00

[a] Cumulative statistic, across all six subsystems

T_{wm}, σ_{wm} - Model 3 event waiting time average, standard deviation (sec.)

T_{we}, σ_{we} - Empirical event waiting time average, standard deviation (sec.)

λ_c, μ_c, k_c - Control task customer arrival rate and service rate (customers/sec.), and Erlang distribution shape parameter

μ_s, k_s - Subsystem event service rate (events/sec.), and Erlang distribution shape parameter

Figure 1. CRT display showing simulated cockpit instruments

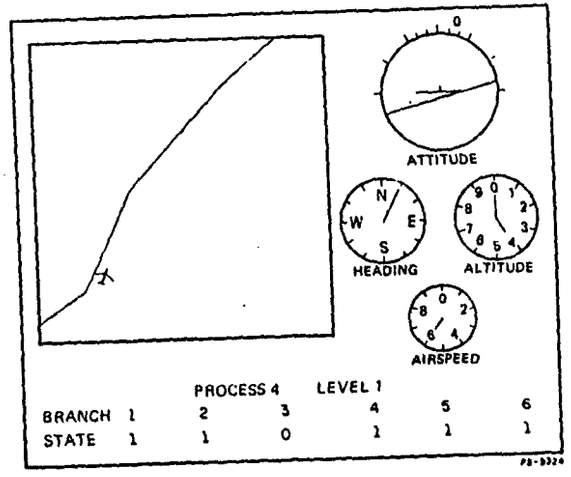


Figure 2. CRT display after pilot response to subsystem event

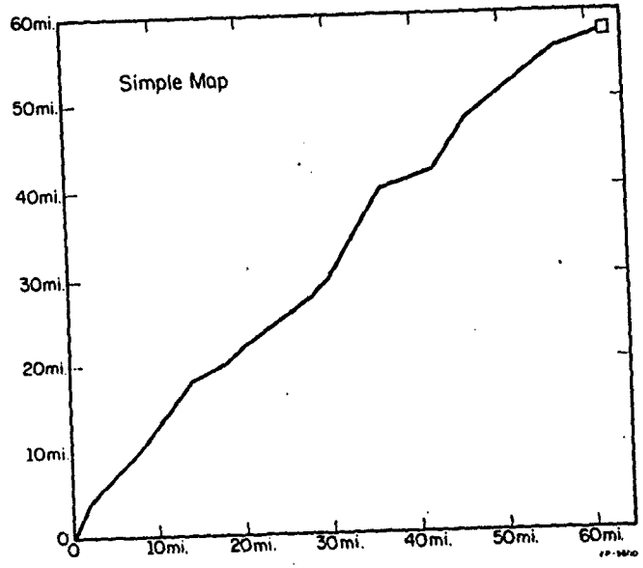


Figure 3. Simple map

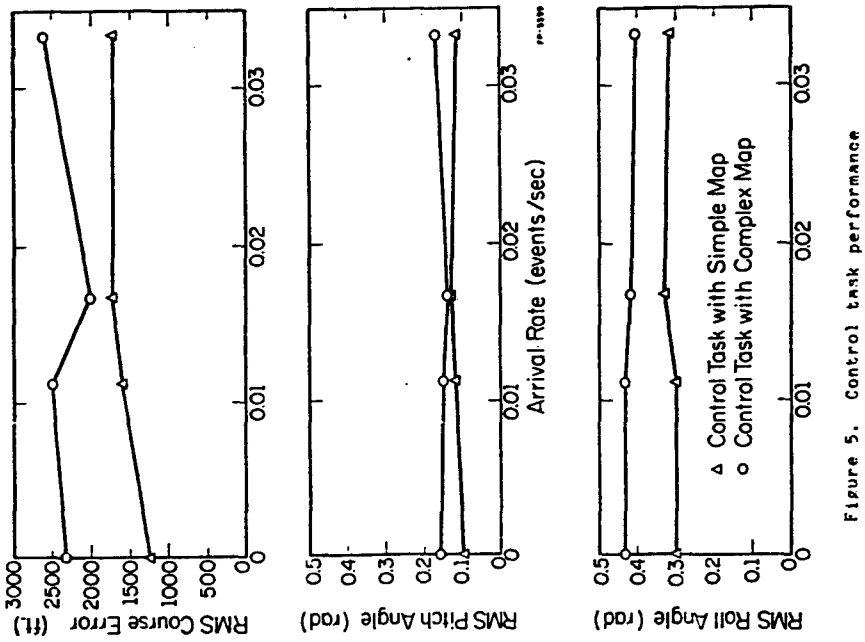


Figure 5. Control task performance

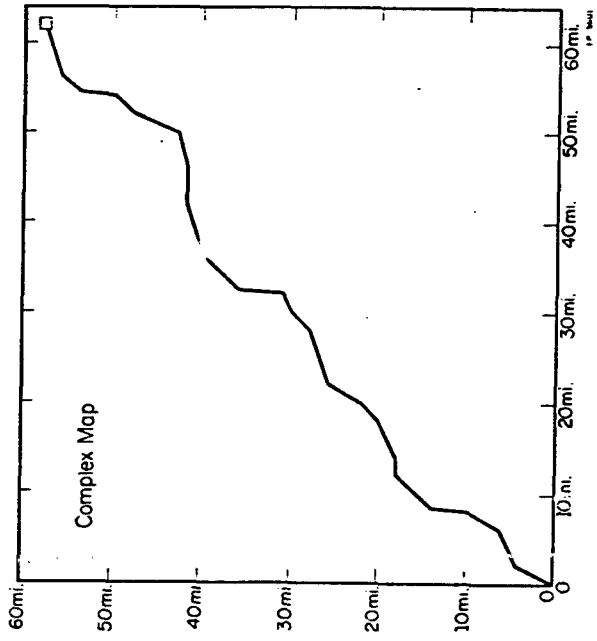


Figure 6. Complex map

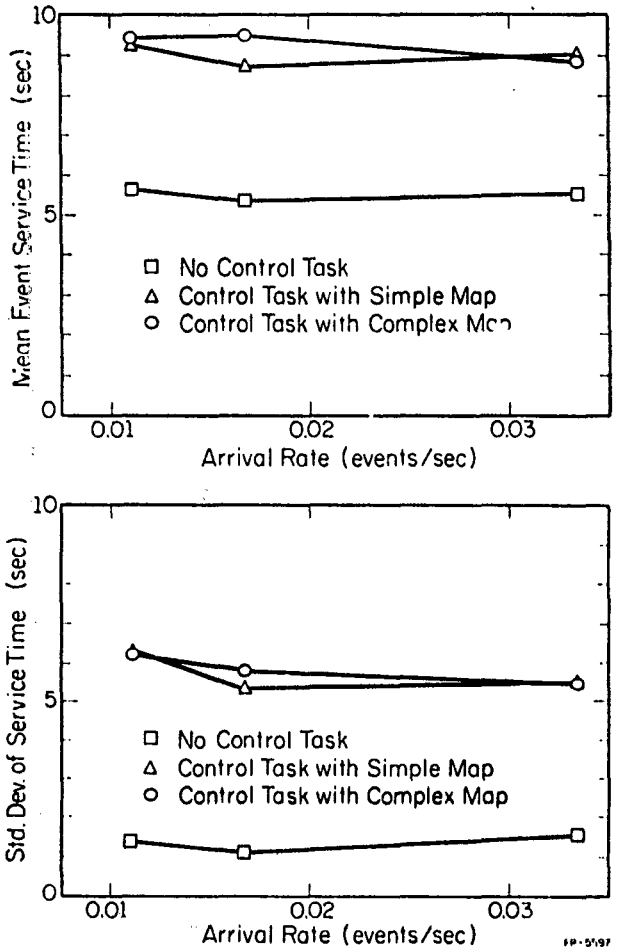


Figure 6. Subsystem event service time statistics (empirical)

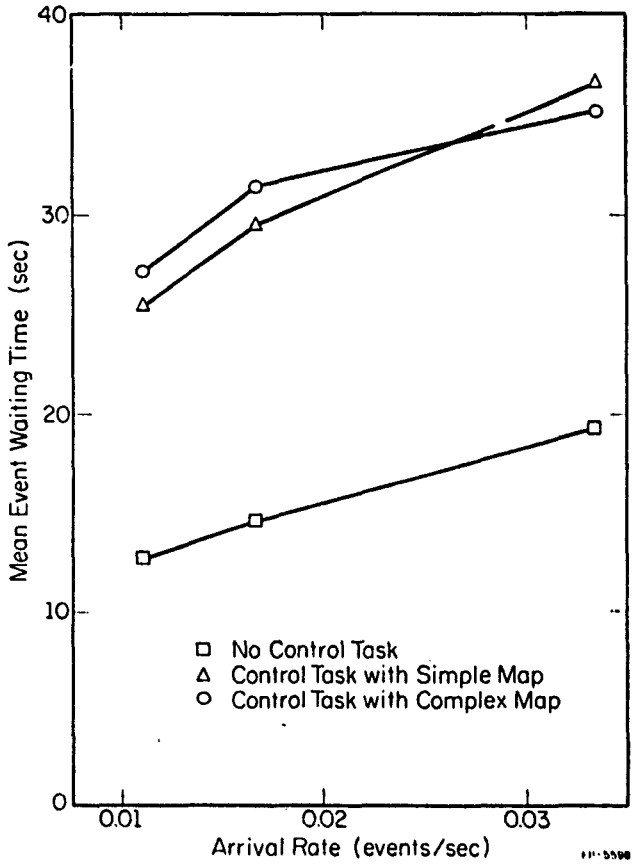


Figure 7. Average subsystem event waiting time (empirical)

Figure 8. Aileron control stick input

N79-17498

INTERRUPTED MONITORING OF A STOCHASTIC PROCESS

by Everett Palmer

Ames Research Center, NASA, Moffett Field, CA. 94035

ABSTRACT

As computers are added to the cockpit, the pilot's job is changing from one of manually flying the aircraft, to one of supervising computers which are doing navigation, guidance and energy management calculations as well as automatically flying the aircraft. In this supervisorial role the pilot must divide his attention between monitoring the aircraft's performance and giving commands to the computer. In this paper, normative strategies are developed for tasks where the pilot must interrupt his monitoring of a stochastic process in order to attend to other duties. Results are given as to how characteristics of the stochastic process and the other tasks affect the optimal strategies. The optimum strategy is also compared to the strategies used by subjects in a pilot experiment.

INTRODUCTION

"New York control, this is NASA 1 arriving on CARMEL 2 with an expected arrival time at MERGE waypoint of 14:31:00." "NASA 1, you are cleared to arrive on CARMEL 1, with a merge time of 14:32:10." This exchange between pilot and controller occured in a recent Ames simulation study of 4D RNAV in the terminal area(1,2). The pilot was cleared for a different RNAV approach route and arrival time. The pilot next entered this data into his onboard navigation and guidance computer. In doing this he had to divide his attention between monitoring the autopilot's performance with his flight instruments and entering data into the computer through his multifunction display and keyboard. Observations of how pilots divided their attention between monitoring and data entry tasks were the motivation for the modeling and more structured experimental work on attention sharing presented in this paper.

The environment in which the pilot interacts with his onboard computer is quite different from other jobs where a person interacts with a computer. In a management information system, teleoperator control, or in most human interaction with a computer, the computer is, or can easily be halted to allow the person time to think and plan his next input. The person and the computer work sequentially. When an aircraft is being controlled in real time by a computer it can not be stopped while the pilot leisurely inputs his commands. In this environment both computer and man must work in parallel. The pilot must interrupt his monitoring to interact with the computer. He must also interruot the discrete tasks to monitor. Other characteristics of discrete tasks and monitoring in the cockpit are the following. The discrete tasks are presented at random. They should be accomplished by a certain time but usually sufficient time is available to do the tasks. Attention must be diverted from monitoring for fairly long blocks of time (seconds) to do the discrete tasks. The displays the pilot must monitor show the error between his vehicle's state and the desired state. When the aircraft is controlled by an autopilot, these signals are relatively low bandwidth signals that should be monitored for out of tolerance readings.

The objective of this research is to determine how design parameters of both displays and the computer interface affect monitoring and data entry performance. In this paper, a task is developed which has many of the above characteristics but which is simple enough so that the attention allocation problem has an optimal solution. A model based on the internal model concept is developed for this task. The model treats the discrete tasks as a constraint and then uses a dynamic programming formulation to maximize monitoring performance subject to the constraint of finishing the discrete tasks on time. Results are presented as to how well pilots could monitor a first order process for out of tolerance signals when they were allowed to monitor the display only half the time.

SPECIFIC PROBLEM

Process Dynamics: The process to be monitored is the output of a first order filter driven by white gaussian noise. The display (fig.1) is updated every 2 seconds and is quantized into 11 cells - .5σ wide. The display is defined as being out of tolerance if it is in the outermost 2 cells ($|x| > 1.75\ \sigma$). In figure 1 these states are indicated with + signs. The process bandwidth determines how predictable the signal is. The ratio of the tolerance to the output variance determines how frequently the signal will be out of tolerance.

Monitoring Task: Whenever the subject observes the process as being out of tolerance he gets a reward of one unit.

Discrete Task: At each time at which the display is observed, the subject decides to either monitor next time or to divert his attention to the discrete task for one or more units of time. The subject was constrained to do m discrete tasks in the next n time units. The objective was to maximize the reward for monitoring subject to the constraint of finishing the discrete tasks. This constraint formulation seems to be a good description of the real situation. It does not require the experimenter to specify a cost functional stating the relative worth of time spent on monitoring and discrete tasks.

THEORY

A review of the literature in the fields of manual control, human factors and psychology found a number of empirical studies which required the operator to interrupt monitoring tasks to do discrete tasks. Models have also been developed for either instrument monitoring or discrete tasks. No papers were found which addressed the problem of what strategies operators use, or should use, to time share their attention between monitoring and discrete tasks. However, Smallwood's paper (3) on human instrument monitoring proposes an approach which can be applied to the present problem. This approach makes the reasonable assumption that the operator has an internal model of the process he is monitoring and of the environmental factors that affect the process. This internal model can be used to predict the future behavior of the process. Smallwood makes the following assumptions that describe how the operator reacts to enviornmental inputs.

> Assumption 1: The human operator bases his state of information about his enviornment upon an internal model of this environment; the model is formed as a result of past perceptions of his environment.
>
> Assumption 2: The human operator behaves optimally with respect to his task and his correct state of information- within his psycho-physical limitations.

The structure of this model is shown in figure 2. The key problems in using this approach are to discover the form of the operator's internal model and the optimal response. If the operator's model of the process is exact and he has no psychophysical limitations the resulting model is normative. Introducing errors in the internal model and psycho-physical limitations such as observation noise, reaction time or faulty memory converts the original normative model into a discriptive model of human behavior.

In the following models, it is assumed that the operator's internal model of the process and enviornmental disturbances is exact. He knows the parameters of the process and can use this knowledge to predict the probability of being in a particular state in k seconds given he knows the current state. For a first order process with bandwidth ω_0, the distribution of the position of the display after last observing the display t seconds ago at position x_0 is a gaussian distribution with

$$\text{mean} \qquad m(t) = x_0 e^{-\omega t}$$

$$\text{variance} \qquad v(t) = \sigma_0^2 (1-e^{-2\omega t})$$

Figure 3 plots these distributions and the probability that the signal will be out of tolerance in the future for various values of x_0.

A decision is made after each monitoring observation of how many stages to devote to discrete tasks. The decision may be to do no discrete tasks in which case the operator continues to monitor.

Define

$P_{ij}(k)$ = the probability that the process will be in state j in k stages, given the process is in state i.

$f_n(m,i)$ = the maximum expected return when the process is observed in state i with n stages to go and m discrete tasks remain to be done.

q = the number of stages devoted to discrete tasks before the next monitoring observation when the process is observed in state i with n stages to go.

C = set up cost, the number of stages wasted when attention is shifted to discrete tasks.

$q^* = q-C$ if $q-C \geq 0$

$q^* = 0$ if $q-C < 0$

then

$$f_n(m,i) = \max_{0 \leq q \leq m} \left[\sum_{\text{jout}} P_{ij}(q) + \sum_{\text{all } j} P_{ij}(q)\, f_{n-q-1}(m-q^*, j) \right]$$

$$f_n(m,i) = 0 \quad \text{if} \quad n=m$$

Figure 4 is a cross plot of two measures of the optimal strategy 4 for different values of process bandwidth. The measures are the fraction of time spent doing discrete tasks or not monitoring the display, f(tasks)=m/n, and the fraction of observed out of tolerance signals to the total number of out of tolerance signals, f(hits)= $\Sigma f_n(m,i)/n$. The figure shows that if the optimal strategy is followed when monitoring a first order process with a bandwidth of .2 rad/stage, 85% of the out of tolerance signals will be observed even if only 50% of the time is spent monitoring. Figure 4 also shows the monitoring performance that would be expected if the pilot could make perfect predictions and his expected performance if he could make no predictions. As the bandwidth of the process decreases and the signal becomes more predictable performance approaches that possible with perfect information.

In many discrete tasks there is the equivalent of a set up cost each time the task is started or restarted after being interrupted. For example, in entering data into a keyboard, some time is lost while the pilot shifts his attention to the keyboard and positions his hands. Figures 5 and 6 show the effect of a discrete task set up cost on sampling strategy and monitoring performance. As the set up cost increases, the best strategy is to look away for longer and longer periods of time when the display is observed near the center. With a set up cost of 2, if the display is observed in the center, the best policy is to complete all of the discrete tasks with no interruption. This is

way the monitoring performance shown in figure 6 for a set up cost of 2 is so close to the performance that is possible when no predictions are made.

Figure 7 shows the sensitivity of monitoring performance to discrete task chunk size - the minimum number of stages which must be spent on discrete tasks. Note that when the minimum chunk size is 5 that the decrement in performance is only large when less than 60% of the time must be spent on discrete tasks. This is because when f(tasks) is greater than 60% the optimum strategy is to look away for more than 5 stages so that a minimum chunk size of 5 is only a minor constraint on performance.

Finally figure 8 shows the sensitivity of monitoring performance to display tolerance or the probability that the display will be out of tolerance.

EXPERIMENT

Results are presented in this section as to how well subjects could monitor a first order process for out of tolerance signals when they were allowed to devote only half of their time to monitoring. The objective of this exploratory study was to determine what strategy subjects used and how it differed from the optimal strategy. To this end, the monitoring task was designed to be identical to the monitoring problem solved by the dynamic programing model. It was not expected that the subjects would behave optimally but a number of features of the optimal strategy seem to be fairly intuitive. Namely it is best to look away longer when the display is farthest from the out of tolerance boundary, look away longer and less frequently with a setup cost, and look away longer if the ratio of discrete tasks to do to the number of stages to go approaches 1.

The subjects monitored the output of a first order filter (ω=.2 rad/stage) driven by white noise. The display, shown in figure 1, was quantized into 11 states, .5σ apart. The display states were numbered from left to right with state 6 in the center. Since the display is completely symetric about state 6, the strategy in states 5&7,4&8,.....,1&11 are identical. In the following discusion we will refer to only the left side of the display. The process was defined as being out of tolerance whenever it was in the outermost two states. The signal was out of tolerance 9.0 percent of the time.

The CRT terminal (figure 1) showed two integers to the left of the monitoring display. The left integer showed the number of stages remaining in which the subject had to finish the discrete tasks. The right integer showed the number of discrete tasks left to do (m). The display was updated every 2 seconds. On each update the number of stages to go (n) was decremented by 1. In order to do a discrete task, the subject pushed the space bar on the keyboard and at the next display update only the two integers were displayed and both were decremented by 1. To switch back to the monitoring display the subject pushed the space bar again.

The subjects were instructed to attempt to observe as many out of tolerance signals as possible subject to the constraint of finishing all of the discrete tasks (m) before the number of stages to go (n) reached zero. On some of the runs a set up cost of 1 stage was added to the discrete task. In this case if the subject switched to discrete tasks for k stages, the number of stages to go was decremented by k and the number of tasks to do was decremented by k-1.

One replication of this experiment consisted of 40 stages. The number of stages to go (n) and the number of discrete tasks to do (m) were initially set to 40 and 20 respectfully. When n reached 0, it was reset to 40 and m was reset to 20 and the next replication began.

Four airline pilots served as subjects in this experiment. Each subject monitored 7 blocks of 6 replications each for a total of 1680 stages (7x6x40). During the first block the subject just monitored the display in order to get a feel for the process dynamics. On the next three blocks there were no discrete tasks to do on the first replication but the subject had to look away from the display for 20 stages on each of the remaining 5 replications. On the last 3 blocks a set up cost of 1 stage was introduced.

Subjects were given a 5 minute rest between each block. Each subject monitored the same random process as the other subjects.

For each replication, performance measures included the number of times the display was out and observed out (a hit), the subjects sampling strategy at each stage, and the optimal strategy at each stage. The above data was also collected for an "autopilot" following the optimal strategy for the same random sequence monitored by the subjects.

At the end of the experiment the subjects completed a questionaire in which they were asked questions about the strategy they used and to rate the experimental tasks on a set of semantic differential scales.

RESULTS

Figure 9 shows the fraction of hits for the four subjects, the autopilot,and the model. The "autopilot" results are the optimum performance for the particular random sequence used in the experiment. The model results are the expected performance for an infinitely long random sequence. As expected no subject did as well as the autopilot, however the subjects monitoring performance with no set up cost was well above the chance value of .5. The set up cost however caused a large decrement in per-

fromance. The subjects could have performed better on the average if at the beginning of each replication they had looked away for 21 stages and completed the discrete tasks all at once.

With a set up cost the subjects should have looked away less often and for more stages than they they actually did. Figure 10 shows the frequency of decisions of various lengths for the subjects and the autopilot. With no setup cost the subjects average decision was 1.8 stages whereas the autopilot's average decision was only slightly longer at 2.1 stages. The main difference between the subject and autopilot was that the subjects looked away for only 1 stage almost twice as often as the autopilot.

With the set up cost both subjects and autopilot increased the average lengths of their decisions to 3.6 and 4.2 resoectively. Note however that the subjects diverted attention to discrete tasks almost twice as often as the autopilot (7.6 vs 4.4). The subject's also made a number of decisions of length 1 which with a set up cost of 1 accomplished nothing.

The optimal strategy is always to look away from the monitoring display for more stages when the display is observed near its center (state 6). The data in Table 1 is similar to that in figure 10 but table 1 also shows the effect of display state on subject and autopilot sampling decisions. The subject's decisions were a strong function of the display state. The subjects almost never looked away from the display when it was out of tolerance and only a few times when it was almost out (state 3). The closer the display was to its center, the longer the subjects looked away - in general agreement with the optimal stategy. Note however that in addition to not looking away quite as long as the autopilot, that on a number of occasions the subjects continued to monitor when the display was near the center (states 5 or 6). The autopilot always looked away in these states if there were any discrete tasks remaining to be done. Some of the subjects decisions to continue monitoring were probably due to the forced pace -real time nature of the task because if the subject failed to push the space bar he would continue to monitor by default. This conjecture is partially reinforced by the pilots description of their strategy. All pilots reported that their strategy was to look away for various numbers of stages when the display was near its center.

Figure 11 plots the number of discrete tasks left to do (m) vs. the number of stages to go (n) for the subjects and the autopilot. It shows whether the discrete tasks were done early or late in the replication. With no set up cost the subjects and the autopilot behaved similiarly- both keeping the ratio of m to n very slightly less than 0.50 and thereby spreading the discrete tasks through out the time availiable. Note however that both autopilot and subjects tended to finish a few stages before the end of the replication. The right hand graph shows that with a set up cost of 1 the autopilot and subjects behavior was quite different. The subjects did not look away long enough and therefore got behind in finishing the discrete tasks. The autopilot on the other hand tended to finish early. This type of behavior prevents the autopilot from becoming trapped if the display starts to go out of tolerance with a few stages left to go. The right hand graph indicates that this rational behavior is not what the subjects intuitively did.

Table 2 sumarizes the results of comparing each decision made by the subjects with the optimal decision. In states 1,2 and 3 the subjects tend to perform optimally. In state 4, the subjects look away for more stages than is optimal - especially when there is a set up cost. In states 5 and 6 , the subjects did not look away long enough. These observations hold with and with out a setup cost. However subjects made considerably more optimal and near optimal decisions with no set up cost.

During the debriefing, 2 subjects said that with zero set up cost their strategy was to look away for 3 times in state 4, 2 times in state 5, and 1 time in state 4, and 0 times otherwise. The other two subjects said that they looked away 2 or 3 times in state 5 or 6 and did not look away otherwise. On the average , these strategies will result in monitoring aproximately 50 % of the time. With the set up cost, the subjects stated that they looked away longer (typically 4 or 5 stages) although one subject said he did not feel he could observe the display often enough to do anything but guess. Subjects stated that the display state was the key thing that influenced their strategy. The number of stages to go and the number of discrete tasks to do had little effect on their strategy unless they were running out of time and then they stated they would look away for a longer number of stages.

Figure 12 shows the average subject ratings on a set of semantic differential scales. These adjectives have been ordered so that the ratings for the task with no set up cost are to the right of the ratings with a set up cost. Considering only scales with a difference of one or or more, the task with a set up cost was demanding, hard to learn, confusing ,surprising, annoying, active, complex, and frustrating.

CONCLUDING REMARKS

In this paper the general problem of time sharing attention between monitoring and other duties has been described and a dynamic programing model for attention sharing was presented. Model performance was presented in terms of the fraction of out of tolerance signals seen as a function of the amount of time spent on non-monitoring duties. The effect of such parameters as process bandwidth and tolerance and discrete task set up cost and chunk size on monitoring performance and the normative time sharing strategies was shown. Future work will extend this model to multiple second order processes and incorporate human limitations such as observation noise and risk aversion.

In the experiment subjects monitored the output of a first

order process for out of tolerance signals. The subjects did not perform as well as an autopilot following the optimal strategy for this task. However their time sharing strategies were a strong function of the display state and a weaker function of the ratio of the number of discrete tasks to do divided by the number of stages to go. With a set up cost , the subjects looked away for fewer but longer amounts of time but the optimal strategy required even longer diversions of attention to discrete tasks. The fact that the subjects did not look away as long as was optimal may be attributable to risk averse behavior.

Future experiments will use a continuous version of the monitoring task used in this experiment with second order dynamics. The effect of the discrete task parameters - set up cost, chunk size, time required, and time available - on monitoring performance will be determinod. Process variables will include bandwidth and the number of displays. The results of this experiment and the modeling work will be used to predict monitoring performance in an experiment in which subjects are required to divide their attention between an actual data entry task and display monitoring.

REFERENCES

1.L.Tobias,"Simulator Evaluation of Terminal Area RNAV Concepts",NASA TN , In Press.

2. E. Palmer,"Pilot's Manual for the 4D Area Navigation and Autopilot Systems in the Flight Management Research Simulator",NASA TMX-73188,October 1975.

3. R.D. Smallwood,"Internal Models and the Human Instrument Monitor", IEEE Trans. on Human Factors in Electronics ,Vol. HFE-8,No. 3, Sept. 1967,pp.181-187.

4.R.A.Howard,"Dynamic Programing and Markov Processes", MIT Press,1960.

Table 1. Subject and autopilot strategy as a function of display state for set up costs of 0 and 1. The data are the average number of decisions for each display state for 600 stages of monitoring when n>2 and n-m>2.

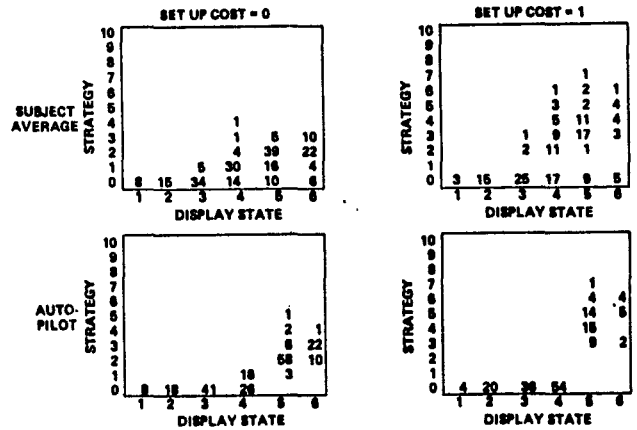
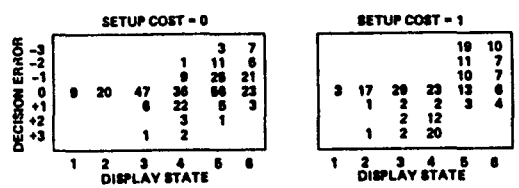


Table 2. Error between the optimum and the subjects decisions as a function of the display state for setup costs of 0 and 1. The data are the average number of decision errors for each display state for 600 stages of monitoring. Decision error = Optimum decision - subject decision.



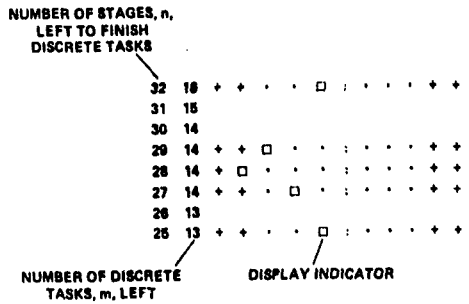


Figure 1. The monitoring display used in the experiment. A new display line was added every 2 seconds. The display was quantized into 11 cells = .5σ wide. The display was out of tolerance if it was in the outermost 2 cells indicated with the + signs. At stages 29 and 27 this subject decided to look away from the display to do discrete tasks for 2 and 1 stage respectively.

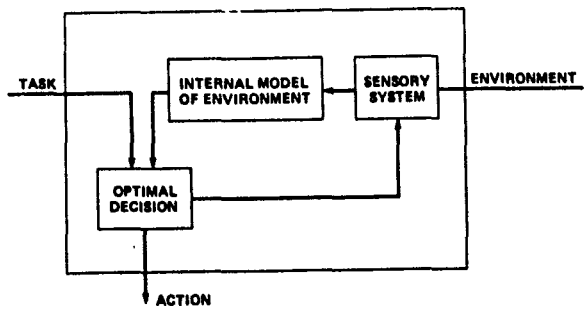


Figure 2. A block diagram of the human monitor (from Smallwood (3)).

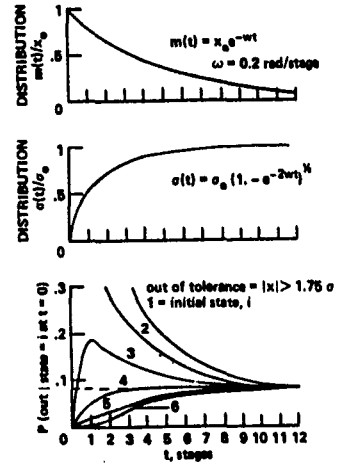


Figure 3. The state of information of a perfect monitor after looking away from the output of a first order filter with bandwidth 0.2 rad/stage driven by white noise.

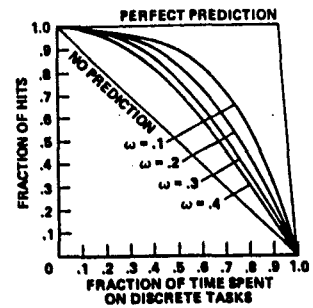


Figure 4. Fraction of hits vs fraction of time on discrete tasks for four values of process bandwidth.
(σ =1.0,T=1.75, P(out)=.080)

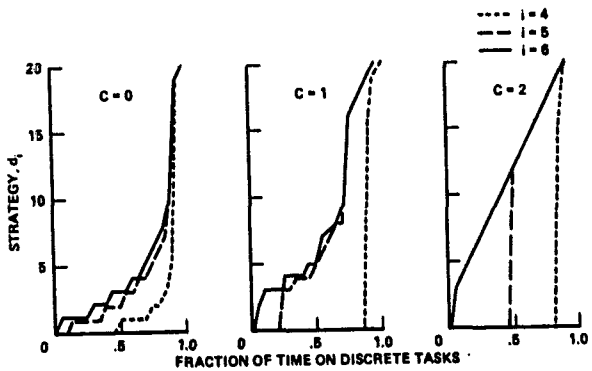


Figure 5. The effect of a discrete task set up cost (C) on the optimal time sharing strategy for states 4,5 and 6. In states 1,2 and 3 the optimal decision is 0 until the fraction of time which must be devoted to discrete tasks is very high.
(n=20 stages, τw =0.2 rad/stage, T=1.75, P(out)=0.080)

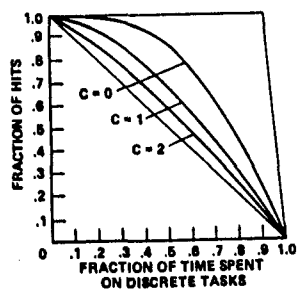


Figure 6. The effect of a discrete task set up cost on monitoring performance for the dynamic programing model with discrete task constraint.
(n=39 stages, τw =0.2 rad/stage, T=1.75, σ =1.0, P(out)=0.080)

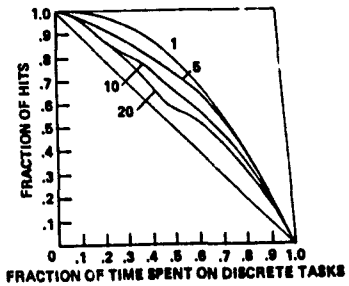


Figure 7. The effect of discrete task chunk size on monitoring performance.
(τw =0.2 rad/stage, n=40 stages, σ =1.0, T=1.75, P(out)=0.080)

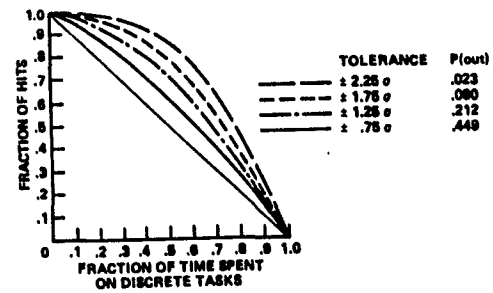


Figure 8. The effect of display tolerance on monitoring performance.
(τw =0.2 rad/stage, σ =1.0, n=40 stages)

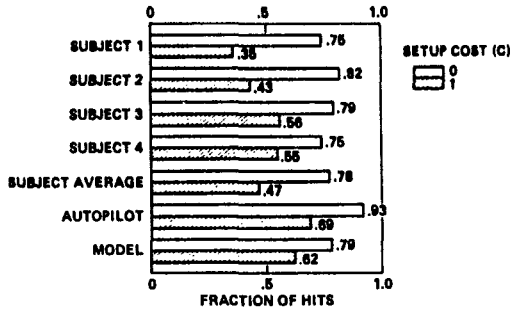


Figure 9. The fraction of hits for discrete task set up costs of 0 and 1.

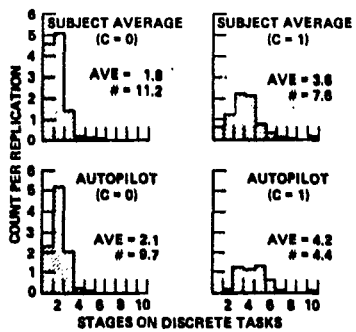


Figure 10. Frequency of decisions to spend various numbers of stages on discrete tasks for the subjects and the autopilot for set up costs of 0 and 1.

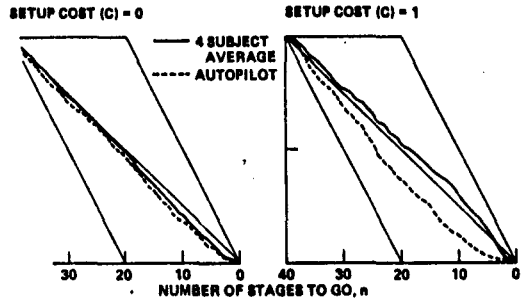


Figure 11. The number of discrete tasks to do (m) vs the number of stages to go (n) for discrete task set up costs of 0 and 1.

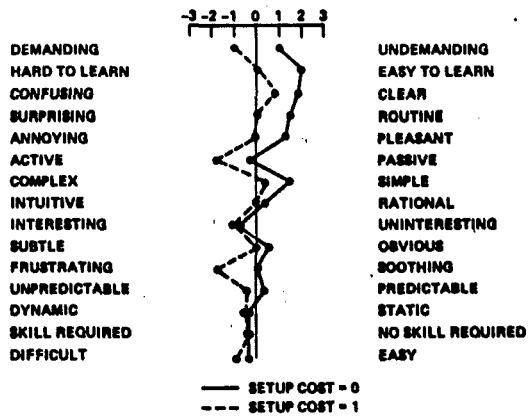


Figure 12. Semantic differential ratings for the experimental task with set up cost of 0 and 1.

NAVIGATOR PERFORMANCE MEASUREMENT IN NAP-OF-THE-EARTH (NOE) MISSION

E. M. Connelly
R. F. Comeau
M. L. Fineberg

ABSTRACT

A Nap-of-the-Earth (NOE) flight path for rotary wing aircraft can be considered as a sequence of terrain features providing visual cues to the navigator/pilot. Salient geographic features along the desired flight path are used as cues (decision points) by the navigator and as initial points and landing zones by the route planner. The successful NOE navigator must use his visual cues to make correct decisions at each decision point along the desired path. Knowledge of a navigator's performance at each type of decision path is required to predict mission success, and to plan and evaluate training programs.

The research reported in this paper is an analysis of NOE flight data to determine the probability of a correct decision at each type of decision point. Also, the effect of flight experience and training on the probability of success at each decision point is identified. A method of coding geographic features termed the Star Code was developed and used to characterize the terrain features at each decision point.

N79-17499

AIR TRAFFIC CONTROL BY DISTRIBUTED MANAGEMENT IN A MLS ENVIRONMENT

J. G. Kreifeldt L. Parkin
Tufts University
Meford, Massachusetts

S. Hart
San Jose State University
San Jose, California

ABSTRACT

The microwave landing system (MLS) is a technically feasible means for increasing runway capacity since it could support curved approaches to a short final. The shorter the final segment of the approach, the wider the variety of speed mixes possible so that theoretically, capacity would ultimately be limited by runway occupance time only.

The dense traffic environment resulting from efficient use of the MLS necessarily reduces the permissible response times and tolerances in the ATC system thus emphasizing the tactical aspects of traffic control.

An experiment using the multi-man ATC facility of the Man-Vehicle Systems Research Division at NASA-ARC contrasted air traffic control in a MLS environment under a centralized form of management and under distributed management which was supported by a traffic situation display in each of the 3 piloted simulators.

Objective flight data, verbal communication and subjective responses were recorded on 18 trial runs lasting about 20 minutes each. The results were in general agreement with previous distributed management research. In particular, distributed management permitted a smaller spread of intercrossing times and both pilots and controllers perceived distributed management as the more "ideal" system in this task.

It is concluded from this and previous research that distributed management offers a viable alternative to centralized management with definite potential for dealing with dense traffic in a safe, orderly and expeditious manner.

INTRODUCTION

In previous papers [1-6] we have discussed some of the advantages and disadvantages that seem to be inherent in operating air traffic control in the terminal area by distributed management as opposed to a centralized ground based ("vectoring") management. These discussions were based on results obtained over a period of 5 years from manned simulation experiments at NASA-ARC in the Man-Vehicle Systems Research Division using the multi-man ATC simulation facility developed there. This facility was designed to capture the absolutely vital pilot-pilot and pilot-controller interactions that are lacking in smaller or totally computerized studies.

The results of each experiment have consistently indicated that distributed management, made possible by the appropriate display of traffic information in the cockpit, exhibited highly desirable characteristics such as a smaller interarrival time variance than did the ground based management alternative as well as dramatically reducing controller verbal workload[3,6]. Thus distributed management has a definite potential for coping with present and projected increases in traffic densities and rising costs in a natural, safe, orderly and expeditious manner without increasing system workloads.

The experiments leading to the above statements are by no means exhaustive. Other possible traffic procedures could be devised to deal with increasing traffic densities even without the concept of distributed management and the natural question arises as to whether or not such procedures could themselves be accommodated under a distributed management regime.

Microwave Landing System.

The microwave landing system (MLS) has been proposed as a technical advance which could support an increase in airport capacity and safety.[7,8,9] The anticipated precision and coverage of the MLS could define precise curved descending approaches thus increasing capacity in existing airport space through two mechanisms. The precision together with curved approaches could permit independent operations to closely spaced runways and closer spacing effectively increasing airport capacity. Together with or independently of multiple runways, the precise curved approaches could also be used to shorten the long common final all A/C must share on present ILS approaches. A long common final, low position precision and wake turbulence necessitates considerable interaircraft spacing between aircraft of different speed and weight characteristics. Reducing the common final approach length also increases runway capacity.

Although considerable technical work is taking place in MLS design for civil aviation, no experimental results have been found investigating the impact of the MLS on air traffic management incorporating the capability for multiple curved approaches as a means for increasing runway capacity.

EXPERIMENTAL OBJECTIVES AND PURPOSES

This experiment had two primary objectives:

A) determine the feasibility of distributed management in a MLS environment of multiple curved descending approaches

B) study the impact of multiple curved descending approaches on air traffic management.

The purposes for each objective were respectively:

A-1) increase the range of understanding of the distributed management concept

A-2) accrue evidence as to its situational robustness

B-1) provide basic information pertinent to air traffic management in the MLS environment

B-2) obtain pilot-controller reactions to MLS capabilities

B-3) provide basic information useful in the design of a MLS terminal traffic area utilizing multiple curved approaches.

Both objectives were accomplished under the same experiment. However, this paper will address objective A only. A second paper[10] addresses objective B.

METHODOLOGY

The methodology used to accomplish the stated objectives was essentially similar to that employed in prior experiments.[1-6] Limited descriptions of common features will be given here. Items peculiar to this experiment will be described in more detail. Reference 10 may also be consulted for further descriptions.

FACILITIES

The basic simulation facility comprises:

- three fixed-base simulators with throttle, aileron and elevator control and CRT flight information and status displays
- two-man controller station with a CRT displayed terminal traffic and an alphanumeric graphic I/O terminal
- a programmable intercommunication system linking all participants
- an SEL 840 computer with a E&S line drawing system.

All flight dynamics etc., for each simulator as well as complete experimental control are performed by the SEL 840. All CRT graphics displays are supported by the SEL 840 and the E&S graphics system.

This facility is being expanded and upgraded as described in another paper[11] but the basic interactive nature of the simulation facility is retained.

TASK

1. Route Structure

The basic task was structured around multiple curved approaches to two closely spaced STOL runways as shown in Figures 1 and 2.

The problem begins about 5 1/2 nm along each approach from threshold. All routes merge at the Final Approach Fix (FAF) 1 nm from threshold. The routes have different turn radii with the Mercury Approach (18L) being most severe. This differs from the technique in Benner, et al[8] in which the approaches tangentially intersected a common curved path but at different "merge" points. Table 1 shows the turn radii and required bank angle at 70 kts for each approach used in this experiment.

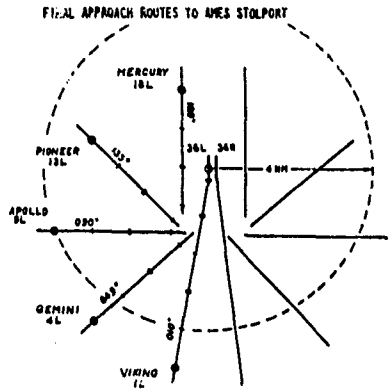


FIGURE 1 FINAL CURVED STOL APPROACH ROUTES TO AMES STOLPORT. SIMULATORS AND COMPUTER CRAFT FLEW APPROACHES TO RUNWAY 36 LEFT. GROUND PROJECTIONS ARE SHOWN.

TABLE 1 - CURVED APPROACH PATHS

Approach	Radius (ft)	Bank Angle (70 kts) deg.	Curved Segment (ft)
MERCURY (18L)	1616	15	4795
PIONEER (13L)	2424	10	5288
APOLLO (9L)	3232	8	4512
GEMINI (4L)	4040	6	2468
VIKING (1L)	-	-	-

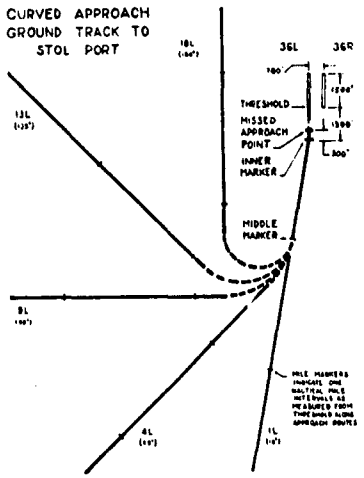


FIGURE 2 DETAILS OF THE CURVED APPROACHES.

Radius and bank angle are monotone functions of the approach route but PIONEER (13L) had the longest curved segment. The range of bank angles is similar to that studied in Benner, et al.[8]

The MERCURY approach (18L) was only used for go arounds. No simulator or computer craft was introduced on 18L.

The curved segments of the ground projection were omitted on the CRT displays because of their considerable drawing overhead. Previous experiments[12] had indicated that path predictors allowed pilots to fly omitted curved segments using their path predictor as a "fairing" device. It was hoped that this effect could be advantageously used in this experiment.

All approaches had the same required altitude profile beginning at 2000 ft, intersecting a 6° glide slope at three miles from the threshold along the approach and descending to 100 ft at the missed approach point (MAP) 1000 ft from threshold. Figure 3 details this profile for reference.

A wind shear of 25 kts at 2000 ft linearly decreasing to 15 kts at ground level blew constantly from the North. No turbulence was present.

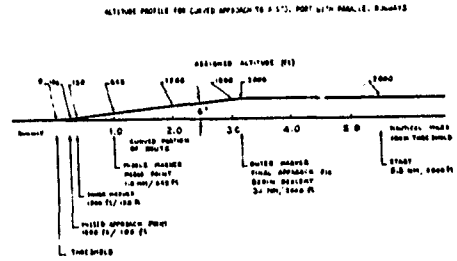


FIGURE 3 ALTITUDE PROFILE FOR EACH OF THE CURVED APPROACHES TO THE SIMULATED AMES STOL PORT.

2. Pilot Equipment

Pilot displays were CRT generated[1,5] with a Vertical Situation Information (VSI) display and a Horizontal Situation Information (HSI) display. The HSI displayed a moving map of ground tracks, A/C symbols and 60 sec. path predictor on own ship only. Previous experiments[2] had determined that path predictors on other A/C were not necessarily an advantage and therefore were eliminated. This also helped reduce drawing requirements. Figure 4 shows a typical pilot CRT display in which other traffic is also visible.

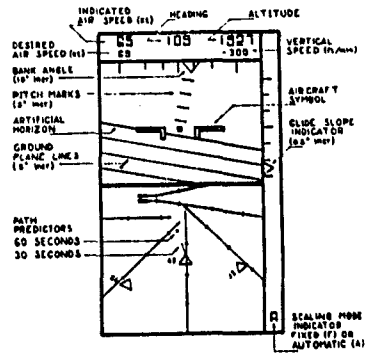


FIGURE 4 TYPICAL CRT COCKPIT DISPLAY. LOWER DISPLAY SHOWS OWN SHIP CENTRALLY LOCATED AND WITH A 60 SEC. PATH PREDICTOR.

A moving pointer glide slope indicator was displayed on the right of the VSI. Horizontal deviations were detected from the HSI which was autoscaled by altitude as an option. Obviously, better indicators should be devised.

Pilots had a switch selectable choice of a translating North up map with a centered rotating A/C symbol or of a heading up centered A/C symbol and rotating translating map.

A second option allowed selection of either a map scale fixed (F) to a constant altitude of 2000 feet encompassing a half width of 12 nm or an automatic scaling (A) option which monotonically decreased the half width with descending altitude to approximately 0.8 nm at 100 ft. The scaling algorithm requires decreasing sensitivity with altitude otherwise vertigo sensations and loss of perspective can easily result as altitude decreases and fluctuates. Automated map scaling requires considerable design thought and experimentation.

Pilots had full use of aileron, elevator and throttle control. Desired Air Speed corresponded to the Indicated Air Speed in straight and level flight which would be achieved at the present throttle setting.

Approach plates for each of the 5 approaches were available in the cockpits.

3. Controller Equipment

The controller station was remotely located from any of the simulators. Two controllers manned the station with one positioned before a CRT showing the Approach ground tracks and the A/C. The second controller sat before a Hazeltine CRT (text display) I/O Terminal which displayed flight number, A/C type, speed, altitude and heading for each A/C. A/C were listed with the most recent flight at the top. Note that the traffic display showed only the A/C flight number. The symbol itself displayed heading information. Altitude and speed information, however, had to be obtained from the text display CRT.

Separating the text from the traffic display A/C symbols was an expedient decision to reduce the amount of graphic display from the E & S unit which, because of the alphanumerics would have been prohibitive.

The controller at the text display could enter selected information via the keyboard. For example, he was responsible for entering A/C speed commands (under the traffic controller's request) to computer generated A/C which were otherwise preprogrammed to fly the approaches along the specified altitude and speed profiles. In this capacity, the text display controller was designated the "super pilot". This role will be discussed in more detail later.

4. Aircraft Types

Both simulators and computer craft were programmed to have simple STOL speed characteristics. The Simulator dynamics were modified NAVION dynamics.(12) Two types of STOL craft were simulated as shown in Table 2.

TABLE 2 - SIMULATED STOL TYPES

Aircraft Type	Speeds (kts) Maximum	Terminal	Approach & Landing	Stall
DHC-7	240	120	65	50
YC-15	425	150	80	60

Computer craft had preprogrammed speed profiles from terminal to approach speed. The super pilot controller could type in commanded speed changes which were accepted (with reasonable inertial response) as long as the resultant speed fell between stall and terminal speeds. Otherwise an error message was immediately shown on the CRT. All computer craft followed their approach route exactly and no path stretching, go around or alternate routes were permissible.

The 3 simulators were only restricted between maximum and stall speeds. The IAS display indicated STALL if such were the case.

An approximately 50-50 mix of the two types of A/C was implemented

5. Task Descriptions

The A/C were required to fly the curved approaches to the 100 ft wide runways spaced 750 ft apart on centerlines. Simulators made approaches to runway 36L only along with other computer craft. Only computer craft made approach to 36R. The MERCURY approach (18L) was reserved for any simulator go arounds.

The basic task required all A/C to cross the 1000 ft missed approach point at 60 sec. intervals on a 6 degree glide slope and at the proper landing speed for the A/C type (65 or 80 kts).

Aircraft were continuously introduced into the problem at 5.5 nm out on an approach at approximately 1 minute intervals. The approach was randomly selected from those not just used on the previous introduction. Any simulator which completed a successful approach was placed into the queue awaiting reintroduction. Thus each simulator could fly several new approaches during the single run of about 20 minutes.

After passing the 1000' marker, a simulator was automatically landed if the glide slope, heading and lateral deviation were within a narrow performance window. If these window requirements were not met, the pilot's CRT flashed GO AROUND and he executed a missed approach to 18L. Controllers could also request a pilot to execute a go around.

Computer craft made perfect approaches and landings (except possibly for speed). No go arounds or missed approaches could be executed with them. This caused considerable problems which will be discussed later and which will be rectified in future experiments.

6. Experimental Conditions: Distributed-vs-Centralized ATC

Two traffic management conditions were studied; A distributed management concept (Sequencing) and a ground centralized procedure (Vectoring) both used in previous experiments. In sequencing, pilots had full traffic information displayed on the CRT HSI, that is they had a Traffic Situation Display. In vectoring, pilots saw only their own A/C symbol.

In sequencing, controllers issued sequence information only to each simulator as it made its approach and refrained from any other statements interpretable as a request or command except for status information. Pilots navigated as needed on their approach to establish their given sequence and satisfy the basic task requirements.

In vectoring, pilots followed controller requests. The super pilot controller in both cases executed the appropriate speed inputs for the computer craft as requested by the traffic controller. The super pilot also provided the traffic controller with flight information if requested from the CRT text display which was also visible to the traffic controller.

In both management conditions, the problem was solved in a man-intensive rather than computer-intensive fashion. Special purpose automated sequencing, spacing and metering algorithms could perhaps have been devised for this particular experiment. However, the history of efforts on that approach suggests that a universally viable procedure should make intensive use of the intrinsic human capabilities already present in the system.

Data were obtained from 3 groups of three professional pilots and two professional controllers per group. Each group made three experimental runs of 20 minutes under each of the management conditions. The number of approaches varied somewhat among groups. Table 3 outlines the experimental design for this study.

TABLE 3 - EXPERIMENTAL DESIGN FOR CURVED APPROACH AND DISTRIBUTED MANAGEMENT STUDY.

		GROUPS: 3 COMMERCIAL AIRLINE PILOTS, 2 AIR TRAFFIC CONTROLLERS		
		GROUP 1	GROUP 2	GROUP 3
TRAFFIC MANAGEMENT SYSTEM	CENTRALIZED AUTHORITY	2 PRACTICE RUNS 3 EXP. RUNS 54 APPROACHES	2 PRACTICE RUNS 3 EXP. RUNS 71 APPROACHES	2 PRACTICE RUNS 3 EXP. RUNS 57 APPROACHES
	DISTRIBUTED AUTHORITY	2 PRACTICE RUNS 3 EXP. RUNS 53 APPROACHES	2 PRACTICE RUNS 3 EXP. RUNS 61 APPROACHES	2 PRACTICE RUNS 3 EXP. RUNS 54 APPROACHES

Each group completed its sessions in one day. The morning was spent in familiarization with the equipment, procedures and task while the experimental session with 3 data runs and a practice run on each management conditions occupied the afternoon.

Subjects were first briefed and given a set of instructions completely describing the experiment. Any questions were answered and discussed as necessary.

Open microphones in the cockpits and push to talk microphones at the controller station connected all 5 participants in a common network. Thus all participants overheard every communication. This is a sensitive condition which requires further study.

It should be pointed out that each A/C symbol on each CRT display carried a simple identifying tag as to its flight number. All simulation were labeled A, B or C followed by a numeral indicating its (re)introduction number. Thus a simulator could be tagged as B1, B2, B3--- etc. If a simulator made a go-around its tag and ID did not change. All computer craft had ID's from D on also followed by their (re)introduction number.

Altitude and speed information on each A/C could be exchanged among pilots through direct address or via the controller.

The mix of A/C types was randomly determined at the beginning of each run with the provision that only two simulators could be of the same type. Each simulator and computer craft then retained its type throughout the run. As a new computer craft was introduced, its type was randomly assigned. It was possible to have upwards of 7 and more A/C making approaches to 36L. All A/C making approaches to 36R were computer craft and proceeded without intervention on the controllers part. It was originally planned to require speed control by the controller of A/C on 36R with the spacing results at the inner marker to be used as a secondary task measure. The problem on 36L proved sufficiently difficult to drop the 36R concern in this experiment. The two runways were thus operated independently.

All pilots including the super pilot (for computer craft) announced their arrival over each approach and from that point proceeded according to the problem development. Simulators could be held at any point but were restricted from holding once on their descent (3 nm from threshold). Computer craft could not be held once past the 3 nm approach marker.

7. Data Recorded

As in our previous experiments, selected flight information from each simulator was recorded at 1 second intervals ("objective" data). All verbal communications were tape recorded. In addition, pilots filled out questionnaires after each successful approach and all subjects filled out other questionnaires after each run. All subjects also completed a final questionnaire after completing the total experiment. Thus objective, subjective and verbal data were available for analysis.

The ground track position of each simulator and computer craft were recorded as part of the objective data and later played back. These reconstructions were video taped in real time and in a speed up (10:1) version for visual analysis.

The objective data provide indications of pilot manual workload as well as spacing information throughout the flights and intercrossing times at the missed approach point and final approach fix (the merge point for the curved approaches).

Verbal data were recorded for analysis of verbal workload in the system and content analysis of the messages.

Subjective data were obtained to study the uniquely human perceptions, suggestions, comments, preferences, etc., necessary for a full evaluation of such complex experiments.

The intercrossing time data and some of the subjective response data will be presented in this paper.

RESULTS

A complete presentation of results will be available in later reports. This paper will discuss intercrossing time (ICT) to contrast (1) distributed-vs-centralized management and (2) differences obtained from piloted simulators-vs-computerized A/C. Limited analysis of subjective data will also be presented. Other results are available elsewhere.[10]

OBJECTIVE DATA ANALYSIS

1. Intercrossing times

The intercrossing times between aircraft at the missed approach point are somewhat complicated by the lack of any control other than speed on the computer craft so that simulator and computer craft results must be treated separately. Accordingly, the four combinations of A/C crossings-simulator precedes simulator (SS), computer precedes simulatro (CS), computer precedes computer (CC) and simulator precedes computer (SC), are analyzed independently. The first two pairings (SS, CS) can be taken as representative of simulator only involvements since the following craft is more likely to make all adjustments relative to the preceding craft. The second two pairings (CC, SC) represent computer craft involvement, by the same reasoning. Certainly SS and CC are most clearly representative of a simulator only study and a computer only study. Besides permitting study of management conditions and curved approaches, the mix of simulators and computer craft also provides insight into vital differences between computer A/C studies and "live" studies will be seen.

Figure 5 shows the cumulative distributions of intercrossing times (ICT) under the sequencing and vectoring conditions without distinction as to the four pairings.

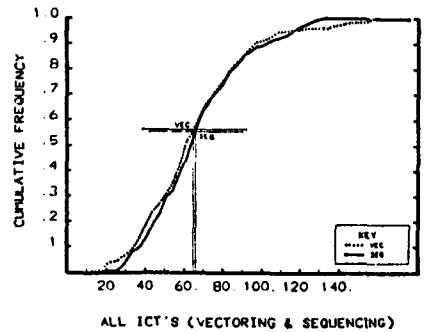


FIGURE 5 INTERCROSSING TIMES FOR THE SEQUENCING AND VECTORING MANAGEMENT CONDITIONS. NO DISTINCTIONS ARE MADE AMONG SIMULATORS AND COMPUTER A/C.

The high degree of similarity under the two management conditions is actually somewhat misleading since results are heavily weighted by computer craft in the sequencing condition where half of the approaches were computer craft handled from the ground as in the vectoring condtion.

Figure 6 shows the ICT distributions for the four pairings under the two management conditions. The differences due to simulator and computer craft behavior are evident here. Note in particular that the smallest ICT values are obtained with computer craft in the vectoring condition which probably represents an inherent lack of responsive control further aggravated by the management condition.

Figure 7 shows the means and standard deviations for the four pairings under sequencing and vectoring management conditions. These are the indicated values from Figure 6. The distinctions between management conditions and computer craft-vs-simulators are easiest to draw from the SS and CC pairings. The mean ICT values are essentially the same at 72 and 70 sec for SS and 61 sec for CC in sequencing and vectoring. Thus there was no appreciable difference in mean ICT between management conditions. The all computer craft crossings (CC), however, show a ICT closer to the specified 60 sec. interval than do the more realistic SS piloted simulators.

However, the standard deviation for SS was about half as large in sequencing (17.7 sec.) as in vectoring (32.1 sec.) while CC again showed negligible differences (23 and 22 sec.).

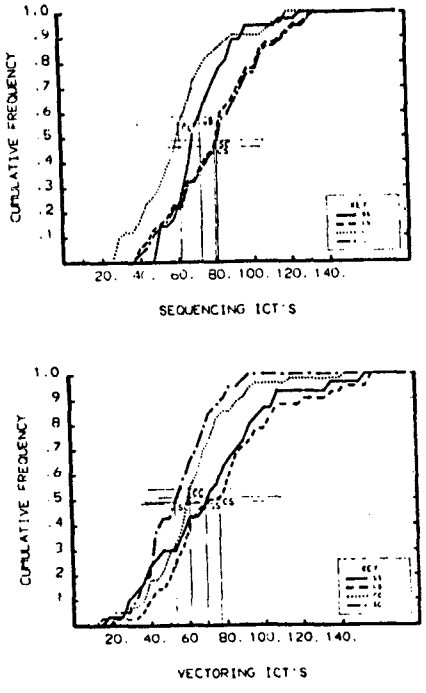


FIGURE 6 SEPARATE ANALYSES OF SIMULATOR AND COMPUTER CRAFT INTERCROSSING TIMES FOR SEQUENCING AND VECTORING MANAGEMENT.

The similarity of means and standard deviations for CC under both sequencing and vectoring is reasonable since the computer craft was always under the super pilot's control and apparently the management condition made no control difference. However, where the pilots could exercise some management, (sequencing) the reduced spread of ICT values demonstrates better system control about the average arrival time.

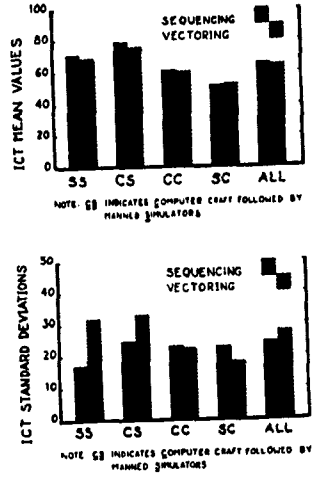


FIGURE 7 DIFFERENCES IN INTERCROSSING TIMES BY MANAGEMENT CONDITIONS AND AIRCRAFT PAIRINGS.

Very similar distinctions can be drawn from the CS and SC pairings in which the standard deviations for CS are about 25 percent smaller for sequencing (24.8 sec.) than for vectoring (32.75 sec.) while SC shows a somewhat smaller spread in vectoring than in sequencing.

Combining the SS with CS crossings and CC with SC crossings provides a broader distinction between simulator active (SS + CS) and computer active (CC + SC) results. Nonparometric statistical analysis showed a nonsignificant difference in mean ICT values for sequencing and vectoring (76 and 74 sec.) but a highly significant difference ($p < .002$) in spreads (22 and 32.5 sec.) between sequencing and vectoring confirming that the pilots in an active role were able to exercise better system control of their ICT than the ground.

The computer active crossings (CC + SC) displayed no significant differences in either mean ICT (57 and 58 sec.) or spreads (23 and 21 sec.) for sequencing or vectoring respectively. Thi is again a reasonable result since the computer craft lacking individual pilots was always directed by the super pilot under the traffic controller's direction.

2. Statistical Analyses

Prior to any statistical analysis an elementary treatment of the data was performed. Because full experimental control was not always available, several non-representative intercrossing times occurred due either to lack of sufficiently responsive control of the computer craft, start up transients, or lulls due to the aircraft introduction statistics. These outliers were removed to arrive at the final number of approaches used in the analysis as shown in Table 4.

TABLE 4 - LIST OF OUTLIER INTERCROSSING TIMES REMOVED FROM ORIGINAL NUMBER OF APPROACHES.

AIRCRAFT PAIRINGS	OUTLIERS REMOVED 20 sec. ICT 175 sec.		CORRECTED NUMBER OF APPROACHES	
	Sequencing	Vectoring	Sequencing	Vectoring
SS	-	-	36	30
CS	192	190	40	42
CC	3,10,15	2,4,8,9,16	43	55
SC	-	11,12,13,14	36	40
TOTALS	4	10	155	167

Some elementary instructive observations can be made at this point. More than twice as many outliers (10) had to be removed under the vectoring condition than under sequencing (4) while the number of approaches was about the same in each case. More outliers had to be removed in the computer craft active pairings (CC + SC = 12) than in the simulator active (SS + SC = 2). Most outliers were removed from the strictly computer active crossings (CC = 8), fewer from a computer craft following a simulator (SC = 4), fewest from a simulator following a computer craft (CS = 2) and none at all from the totally simulator active crossings (SS = 0). This is strictly in accordance with the flexibility of system control due both to distributed management and to live-vs-computer simulation.

Initial x^2 tests of the ICT data did not support a normality assumption for the data in every pairing and management condition as might be supposed by inspection of Figure 6. The normality assumption was rejected for the CC pairing alone in vectoring and for all but CS in sequencing. Rather than depend on robustness arguments or comparison specific tests, a single nonparametric test philosophy[13] was used for all the ICT comparisons.

Essentially, the nonparametric test hypothesizes for each comparison n- tuple that the groups in question are simply randomly selected from the pooled comparison data according to each group sample size; this hypothesis is exercised by computer until stable estimates are obtained for each group distribution and then confidence limits are set. The significance of the obtained statistic (e.g. mean or standard deviation) is then accepted or rejected accordingly. Comparisons between mean values were based on the unsigned difference statistic while standard deviation were compared by the familiar ratio statistic. Figure 8 shows the computer generated hypothetical "no difference" statistics for a mean and a standard deviation comparison. The actual test comparisons are also shown.

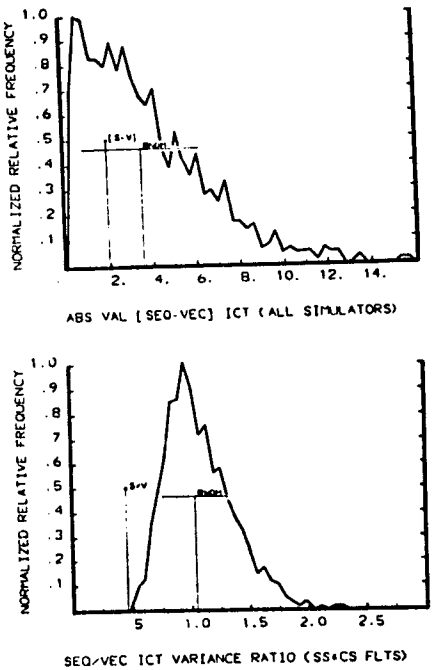


FIGURE 8. COMPUTER GENERATED TEST STATISTICS FOR A MEAN VALUE DIFFERENCES AND A STANDARD DEVIATION RATIO. ACTUAL TEST VALUES ALSO SHOWN.

The significance levels of ICT means and stand deviations are graphed in Figure 9. All the two way comparisons are made between the 4 aircraft pairings.

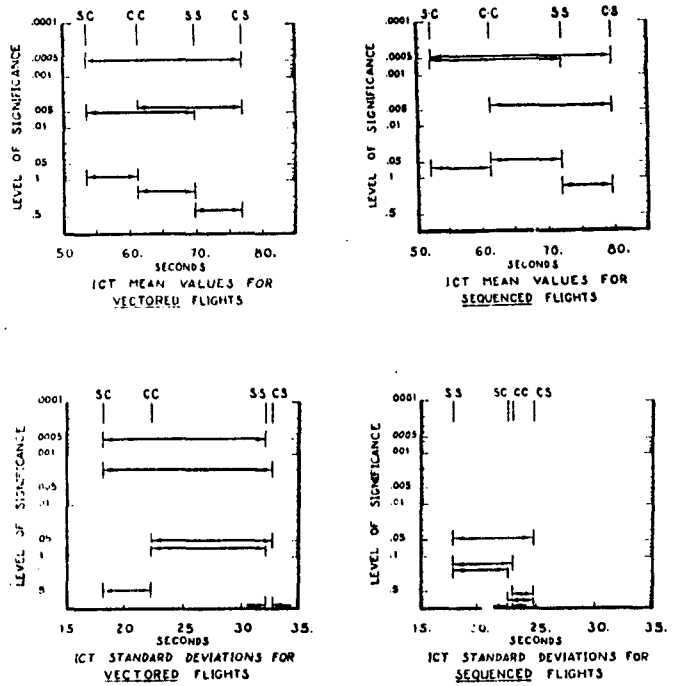


FIGURE 9 PAIRWISE COMPARISONS OF ICT MEANS AND STANDARD DEVIATIONS FOR THE FOUR A/C PAIRINGS. DIFFERENCES BETWEEN THE TWO MANAGEMENT CONDITIONS CAN BE INFERRED.

3. Simulator - Computer A/C Differences

The analysis of effects due to mixing simulators with computer generated A/C is examined in this section.

a. Vectoring

i) Mean ICT

Within the vectoring condition SS and CC produced mean ICT values not significantly different from the grand average of 64.8 sec. although CS and SC with means of 77.0 and 53.3 sec. were significantly different from the grand average at the 0.002 level or better. As Figure 9 shows, there was no significant difference between the all simulator and the all computer pairings (SS-CC) nor between the two simulator active pairings (SS-CS). The two computer pairings (CC-SC) are marginally significant. However, where the comparisons involve simulator-vs-computer active (or vice versa) pairs, the differences are highly significant.

These tests suggest that mixing computer and simulator craft in the same vectoring type experiment can produce very different mean value results due to sequential effects (SC or CS) whereas the homogeneous pairings (SS-CS), (CC-SC) produce more consistent mean values. The smallest ICT (53.3 sec.) was obtained with a computer following a simulator (SC) while a simulator following a computer (CS) produced the longest ICT (77.0 sec.).

ii) Standard Deviations of ICT

The SS ICT spread is significantly larger than those of CC and SC (the computer active cases) in vectoring. Paralleling the results for the mean values, the spread of ICT values is significantly smaller for SC than for CS. No other comparisons showed any significant differences.

These results suggest that vectoring simulators as opposed to computer craft results in greater standard deviations which is in keeping with the mean ICT values.

b. Sequencing

i) Mean ICT

Within the sequencing management condition, the CC crossings had a ICT mean value significantly lower than the SS crossings (61.0 vs 72.0 sec.). In fact, the only nonsignificant differences (at the 0.05 level) were between the pairs SS and CS (both simulator active) and CC and SC (both computer active). This indicates that the pilots received no differential advantage to following another simulator or

a computer craft (ground controlled). The ground likewise received no clear advantage to scheduling a computer craft after either another computer craft or a simulator.

ii) Standard Deviations of ICT

The only standard deviations which appear to be significantly different are those of SS and CS, the two simulator active pairings in which the all simulator case has about half the spread of the mixed case (17.7-vs-24.8 sec.). The SS ICT spread is the only one significantly smaller than the other three.

Combining these observations in the sequencing case suggests that in fact the mean ICT value is closer to the desired value of 60 sec. for the strictly computer A/C (controller managed) than for the strictly simulator A/C (pilot managed) and that the consistency (spread) of ICT values is the same in both cases. There is some indication that the presence of computer craft made things harder for the pilots since the SS spread is smaller than for CS and the mean ICT is better for SS than for CS although this is marginally significant.

In the vectoring case (controller managed all A/C), the mean value is again closer to the 60 sec. value for CC than for SS although the spread was considerably greater when managing the "real" A/C (32.1-vs-22.3 sec.). For some reason it appeared easier to schedule a computer craft to follow a simulator than vice versa and rather than to follow a computer craft or another simulator.

These observations suggest that results based on computer studies not using simulators or using computer-simulator mixes must be cautiously analyzed and interpreted for extrapolation to all simulator behavior. The problem obviously is that nothing substitutes for the inherent human characteristics which, however, subtle they may be, have a profound effect.

4. Management Conditions

Differences due to management conditions (sequence or vector) are examined in this section as reflected in intercrossing times (ICT). Figures 8 and 9 may be reexamined with the results given in Table 5.

The all simulator pairing (SS) may be taken to be the most representative of the results which would arise from an all simulator environment. There was a nonsignificant difference (2.27 sec.) in mean ICT values between sequencing and vectoring producing an average ICT of about 70 sec. However, the standard deviation of ICT in sequencing (17.7 sec.) is nearly 100% smaller than that for vectoring which is highly significant ($p = .001$) both practically and statistically.

Nearly the same observations can be made for the CS pairing in which a simulator followed a computer craft in. There was no significant difference in mean

TABLE 5. COMPARISON OF MANAGEMENT CONDITIONS BY ICT SIGNIFICANCE LEVELS GIVEN.

MANAGEMENT	SIMULATOR ACTIVE PAIRS						COMPUTER ACTIVE PAIRS		
	SS		CS		SS & CS		CC & SC		
	MEAN VALUE COMPARISONS								
	$\bar{X}$	P	$\bar{X}$	P	$\bar{X}$	P	$\bar{X}$	P	$\bar{X}$ = MEAN (SEC)
SEQUENCING	71.97	NS	79.40	NS	75.88	NS	56.48	NS	P = LEVEL OF SIGNIFICANCE
VECTORING	69.70	NS	76.98	NS	73.94	NS	57.83	NS	
SEQ-VEC	2.27	NS	2.42	NS	1.94	NS	1.36	NS	
	STANDARD DEVIATION COMPARISONS								
	S	P	S	P	S	P	S	P	S = STD. DEV. (SEC)
SEQUENCING	17.67	.004	24.83	.086	21.92	.001	23.20	NS	P = LEVEL OF SIGNIFICANCE
VECTORING	32.13	.004	32.75	.056	32.46	.001	20.87	NS	
SEQ/VEC	.550	.001	.758	.080	.675	.0005	1.11	NS	

ICT between sequencing and vectoring but (marginally) significant difference in standard deviations with sequencing producing a spread of ICT values about 50 % smaller than that from vectoring. The reduction in the relative magnitudes of spreads for CS compared with SS must be due to the reduced control flexibility of a computer craft-simulator pair. In the SS pairing the lead A/C can also actively participate in maintaining spacing from the following A/C in sequencing. Removing this opportunity (CS) decreases the control consistency. Notice that in vectoring the spreads of SS and CS were the same showing that the pairing had no effect on the spread.

Combining the simulator active pairs (SS + CS) also produces nonsignificant mean differences between sequencing and vectoring and a highly significant ratio of standard deviations reflecting the previous statements for SS and CS.

The results for the computer active pairings (CC + SC) are included for comparison. Note that the two management conditions do not produce significantly different means nor standard deviations for the computer active pairs. As remarked before the "super pilot" controller handled these A/C the same way ("vectored") in both conditions.

It is obvious that inferences drawn from the CC or (CC + SC) pairings would be very misleading compared to the more realistic SS pairings.

Some of the more important findings based on intercrossing times are summarized.

1. Distributed-vs-Centralized Management.
   - No significant differences in mean ICT(71 sec.).
   - Spread of ICT twice as large in centralized management as in distributed management.

- Shortest ICT (15 sec.) was produced under centralized management. Shortest ICT for distributed management was 47 sec.

These statements are based on simulator-simulator pairings.

2. Piloted Simulator-vs-computerized A/C.

- Simulator mean ICT was greater than for computer A/C in distributed management and nonsignificantly greater in centralized management.
- Simulator ICT spread was nonsignificantly smaller than for computer A/C in distributed management and significantly greater in centralized management.
- Mean ICT was significantly greater for a simulator following a computer A/C than vice versa for both management conditions.
- ICT spread was significantly greater for a simulator following a computer A/C than vice versa in centralized management and nonsignificantly greater in distributed management.
- Largest Mean ICT and ICT spread occurred for simulators following computer A/C. Smallest mean ICT occurred with computer A/C following simulators. Smallest ICT spreads in distributed management occurred for simulators followed by simulators. Smallest ICT spread in centralized management occurred for computer A/C following simulators.

SUBJECTIVE RESPONSES

At the end of the experiment, all subjects were requested to rate the two management conditions on 8 scaled lines reflecting their own feelings. On a scale such as "visual workload", the subject placed a labelled mark somewhere between the LOW and HIGH end. The mark was labeled either vectoring or sequencing. The data were analyzed both for absolute strength of each rating (0-10) and for simple rank order (dominance) of the two conditions. The results are shown in Figure 10.

The pilots felt safety to be somewhat higher in vectoring than in sequencing where as the controllers were fairly unanimous and determined that sequencing was safer than vectoring. This also applies to expeditiousness and orderliness.

Workload particularly visual workload, was judged better in vectoring by both pilots and controllers. Pilots were unanimous in their agreement that vectoring produced a somewhat lower visual workload and total workload than sequencing.

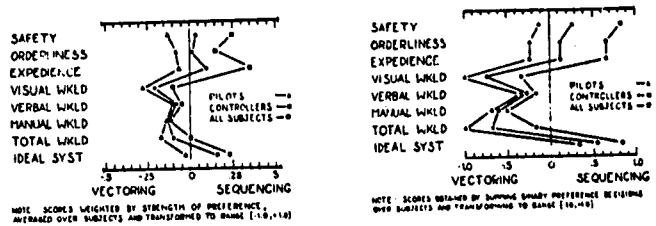


FIGURE 10 SUBJECTIVE RESPONSES BY PILOTS AND CONTROLLERS TO THE DISTRIBUTED AND CENTRALIZED MANAGEMENT CONDITIONS. ABSOLUTE DIFFERENCES ARE SHOWN ON THE LEFT WHILE STRICT DOMINANCE IS SHOWN ON THE RIGHT.

On the average, both controllers and pilots ranked sequencing over vectoring as being closer to their notion of an ideal ATC System for this task. Individually, some pilots dissented from this view vigorously enough to make the average preference strength slightly favor vectoring.

Subjective responses in this experiment presented several anomalies compared to those obtained in previous experiments.[2,5] Previously, controllers felt vectoring to produce greater safety, orderliness and expeditiousness while pilots favored sequencing on these attributes. The present reversals of this stance suggest that the MLS curved approach environment is an entirely new and different task in which neither the pilot or controller group felt itself to have produced the best performance. This may change as the curved approach regime becomes more familiar and/or much higher levels of proficiencies are developed for this task.

Pilots and controllers both felt vectoring to have a lower verbal workload than sequencing which again is a reversal of past results. The verbal analysis presently under way will shed definitive light on this.

However, in spite of the pilots stance favoring vectoring on the first three attributes and its perceived favorable workload measures from both pilots and controllers, the group consensus indicated that distributed management is closer to the notion of an ideal system than centralized management. This apparent contradiction suggests that there are important attributes not included in the ones listed which favor distributed management. It could also be that the "ideal system" concept expresses the pilots' desire to participate more in their own local traffic management (as in previous experiments) and the controllers' desire in this task to assume less responsibility (unlike previous experiments), in favor of the improved performance which occurred under distributed management.

CONCLUSIONS

The following conclusions are drawn from the results obtained in this and previous experiments.

- Distributed management which allows pilots to exercise some tactical control will reduce the standard deviation of interarrival times significantly compared to centralized management.
- Mean intercrossing time will be about the same for distributed or centralized management.
- Studies using directed computer A/C instead of more realistic piloted simulators may provide highly misleading results.
- Distributed management produces system performance at least as good or better then centralized management.
- Distributed management is a task robust concept.
- Pilots favor distributed management. Controllers may favor it depending on the task.
- Traffic management of multiple curved approaches is a complex problem requiring specialized techniques and aids.

In general, distributed management in which controllers assign a sequence to each A/C and allow the individual pilots the freedom to accomplish it is a natural mode of control and produces safe, accurate and consistent results in an orderly and expeditious manner.

REFERENCES

1. Kreifeldt, J.G., Wempe, T., "Future Terminal Air Traffic Management Concepts". Proc. 10th Annual Conference on Manual Control, WPAFB, April 9-11, 1974.

2. Kreifeldt, J.G., Wempe, T., "Human Decision Making in Future ATC Systems: Comparative Studies in Distributed Traffic Management". Proc. 1974 International Conference on Systems, Man and Cybernetics. IEEE-SMC 74 CHO-4SMC Oct. 2-4, 1974.

3. Kreifeldt, J.G., Pardo, B., Wempe, T., Huff, E., "Verbal Workload in Distributed Air Traffic Management". Proc. 11th Annual Conference on Manual Control, NASA ARC, TMX-62, 40.4, May 21-23, 1975.

4. Kreifeldt, J.G., Parkin, L., "Subjective Evaluation with FAA Criteria - A Multidimensional Scaling Approach". Proc. 11th Annual Conference on Manual Control. NASA-ARC. TMX 62,464, May 21-23, 1975.

5. Kreifeldt, J.G., Wempe, T., "Implications of a Mixture of Aircraft With and Without Traffic Situation Displays for Air Traffic Management". Proc. 12th Annual Conference on Manual Control, Urbana-Champaign, Ill. NASA TMX-73, 170, May 25-27, 1976.

6. Pardo, B., Effects on Verbal Communication and Task Variables of Different Air Traffic Control Configurations. Unpublished M.S. Thesis. San Jose State University, August 1975.

7. Cherry, G.W., DeWolf, B., MacKinnon, D., "Increasing Airport Capacity and Terminal Area Safety by Means of the Scanning Beam Instrument Landing System". AIAA Guidance, Control and Flight Mechanics Conference, AIAA No. 70-1033, August 17-19, 1970.

8. Benner, M.S., Sawyer, R.H., McLaughlin, M.D., "A Fixed-Base Simulation Study of Two Stol Aircraft Flying Curved, Descending Instrument Approach Paths", NASA TND-7298, Oct. 1973.

9. Brady, F.B., "Landing Guidance Systems" in A Survey of Modern Air Traffic Control, Vol. II. AGARD-AG-209-Vol. II, 1975.

10. Hart, S., McPherson, D., Kreifeldt, J., "Multiple Curved Descending Approaches and the Air Traffic Control Problem", paper presented at the 13th Annual Conference on Manual Control, MIT, June 15-17, 1977.

11. Kreifeldt, J.G., "Design Outline for a New Multiman ATC Simulation Facility at NASA-ARC", paper presented at the 13th Annual Conference on Manual Control, MIT, June 15-17, 1977.

12. Kreifeldt, J.G., Wempe, T., "Pilot Performance During a Simulated Instrument Procedure Turn With and Without a Prediction Display", NASA TMX-62,201, Jan. 1973.

13. Bock, R. D., Multivariate Statistical Methods in Behavioral Research. McGraw-Hill, 1975.

ACKNOWLEDGEMENT

This work was supported by funds from NASA Grant NSG 2156. The invaluable aid provided by staff and contractors at NASA-ARC is very gratefully acknowledged.

INTERFACE DESIGN IN THE PROCESS INDUSTRIES

M. C. Beaverstock

H. G. Stassen

R. A. Williamson

Every operator runs his plant in accord with his own mental model of the process. In this sense, one characteristic of an ideal man-machine interface is that it be in harmony with that model.

With this theme in mind, the paper first reviews the functions of the process operator and compares them with human operators involved in control situations previously studied outside of the industrial environment (pilots, air traffic controllers, helmsmen, etc.). A brief history of the operator interface in the process industry and the traditional methodology employed in its design is then presented. Finally, a much more fundamental approach utilizing a model definition of the human operator's behavior is presented.

This entire topic is important to the fluid process industries. Indeed, it may be stated that: "He who specifies the process control system (and the accompaning displays) specifies the plant operating procedures - and, hence, the operating economics.

INTERFACE DESIGN IN THE PROCESS INDUSTRIES

M. C. Beaverstock
The Foxboro Company
Foxboro, Massachusetts
H. G. Stassen*
Delft University of Technology
Delft, The Netherlands
R. A. Williamson**
The Foxboro Company
Foxboro, Massachusetts

I INTRODUCTION

The application of control and information processing systems involving complex man-machine systems is still expanding as a result of ever increasing levels of automation. As a consequence, a stage of automation has been reached in which a number of the human operator's mental processes are being taken over or augmented by such systems, resulting in modifications of the human operator's task. Criteria and guidelines for allocation of system functions between man and machine should therefore be established, but a general methodology is presently not available.

The supervisory contol situation can be explained best by the block diagram of Figure 1. A plant is automatically controlled by a set of controllers; the

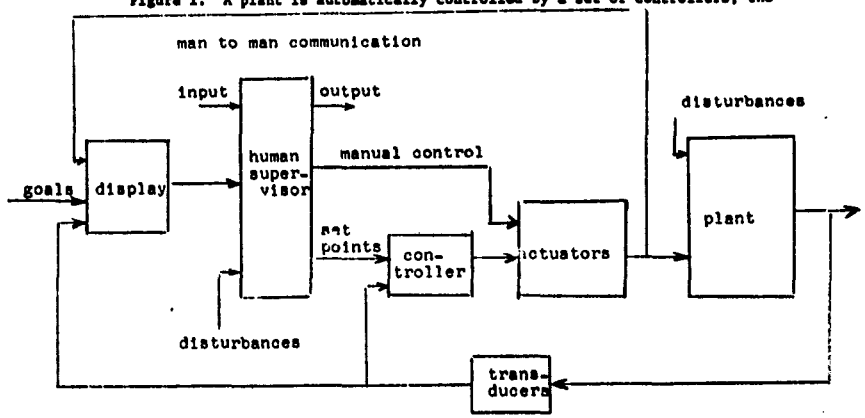


FIGURE 1: SUPERVISORY CONTROL SYSTEM

*Presently he is on sabatical leave at MIT, Cambridge, Mass. for the academic year 1976-1977.

**Presently with Matromation, Inc., Princeton, N.J.

outputs of the transducers are compared with setpoints and revise controller outputs to the final actuators. The human supervisor is informed about the plant conditions via a display system; he monitors the system, he generates setpoints, he makes decisions whether he should intervene or not, he plans control strategies, and he communicates with other human supervisors. The task of the human supervisors can be summarized as follows (Figure 2):

- Learning, understanding, and interpreting goals required of him (actually constructing).
- Monitoring the system outputs so that from the control actions, he can identify the dynamics of the system and the noises acting on the process.
- Planning and determining which control actions should be performed.
- Inputing the appropriate data to the control system for both initialization and on-line adjustments.
- Intervening in order to switch from manual to supervisory control.

The tasks mentioned above deal with three categories of system functions (Rijnsdorp, Rouse, 1977):

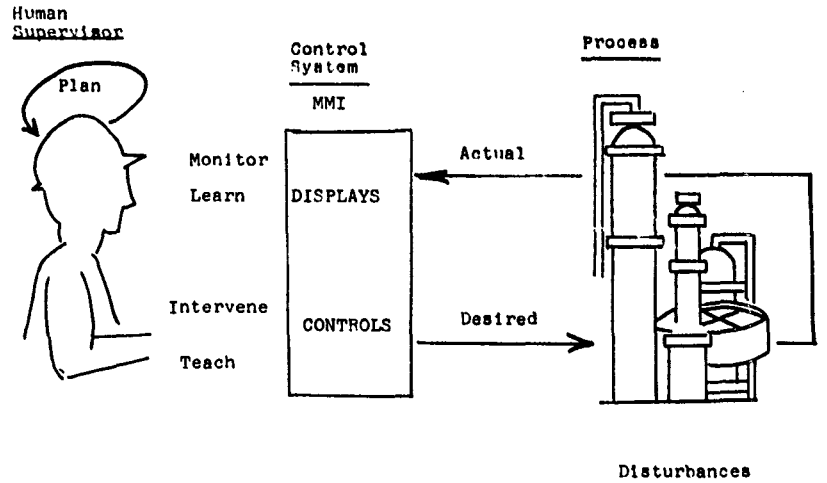
- Supervision and control of process operation.
- Coping with malfunctions.
- System evaluation and improvement.

The human operator has been studied in terms of his tasks and function in many situations (pilot, air traffic control, etc.) other than an industrial environment. In the fluid process industries (chemical, petrol, textiles, food and drugs, metals, pulp and paper), the human supervisor's behavior (and his task) is quite different from that in the many reported studies in vehicle control.

Four different aspects of human supervisor data processing can be recognized with regards to the information to be handled:

- INFORMATION: what type of information should be presented and how much of it?
- PRESENTATION: how should the information be presented?
- MANIPULATION: how should the information be manipulated by the operator?
- CONFIGURATION: how should the presentation and manipulation characteristics be physically arranged to form a work station for the human operator?

With reference to these four aspects we can distinquish the differences between vehicle control and process control; these differences are listed in Table 1. In addition to this table we should add one other important fact. Usually a plant is designed initially from a process point of view, then controls are considered, and finally the human tasks are determined; an integrated design of the man-machine function is seldom employed.



TASKS

LEARN	⟷	Build Internal Model
MONITOR	⟷	Adjust Internal Model
INTRRVENR	⟷	Assume Manual Control
TEACH	⟷	Generate Instructions: Set Points
PLAN	⟷	Determine Control Strategy

FIGURE 2: THE TASK OF THE HUMAN SUPERVISOR

However, displays for these operators in the fluid process industries as well as the automatic control systems, play a very critical role in the operation of a whole plant. Indeed, it may be stated that: "He who specified the process control system (and the accompanying displays) specifies the plant operating procedures - and hence the operating economics." This means that the designer also determines the task of the supervisor and really allocates the system functions between man and machine. Therefore the designer should not only be knowledgeable about the control possibilities and process dynamics of the plant, but he also should be able to take into account human capabilities.

Data Processing Aspects	Vehicle Control	Process Control
INFORMATION		
Type	Direct	Indicative
Dimension	About 10	Between 100 and 1500
Time constants	From sec. to min.	From min. to hours
Abstractness	Low	High
Origin disturbances	Known	Unknown
System properties	Linear	Non-Linear
	Constant	Time-varying
PRESENTATION		
Integration	Integrated displays	Non-integrated displays
Parallel/series	Mostly parallels	Parallel and/or series
Number of displays	Low	High
Time window	Instantaneous and predictive	Instantaneous and history (trend)
Modes	Visual, auditive, vestibular	Visual, auditive
Overview	Good	Poor
MANIPULATION		
Direct feedback	Proprioceptive	No proprioceptive
Response	Immediate	Next shift
Strategies	Error correction	Boundary control
	More or less consistent	Variable
Accuracy	Quantitative	More or less qualitative
CONFIGURATION		
Physical size	Small	Large
Personnel training	Highly and consistently	Less rigorously
Human factors	Well developed	Very little developed

TABLE 1: COMPARISON BETWEEN VEHICLE CONTROL AND PROCESS CONTROL

Even with such incentive, little has been done to systematically study the way an operator carries out his assigned task to operate a plant and then make that knowledge available to a designer to define the appropriate display. To understand this solution requires a realization of the interfaces between man and the plant he controls as it exists in the process industries. Over the years, while plant operating practices have hardly changed, the interface has gone through significant changes so that before thinking about the future it is worthwhile to consider the past.

II BRIEF HISTORY OF THE HUMAN OPERATOR'S FUNCTION

Looking back one sees that the operator/process system was tightly coupled. Measurement and control devices were located next to, or on, the associated equipment. The tools an operator worked with for viewing and manipulation were such things as gauges, handwheels, levers and sight glasses. The operator was "in touch" with his environment as shown in Figure 3. The operator himself determined the operating procedure. His "experience" was a key to the successful operation of the plant. He constantly came into physical contact with the process. His sense of smell could detect overcharging of an ingredient or an incomplete reaction; by sight and touch he could tell the quality of the material he produced; and a slight vibration could warn him of an imminent pump failure. However, the process equipment tended to be widely separated and, therefore, much manpower was required and expended, and control in general was much less than perfect.

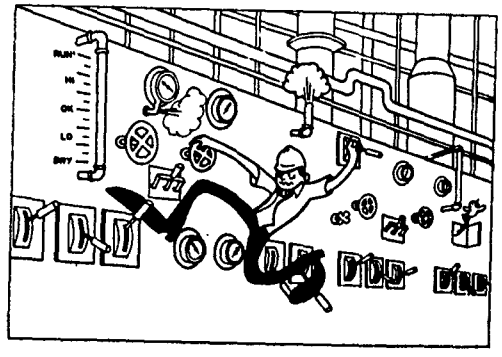


FIGURE 3 - Operator "In Touch" With The Process

As the size and complexity of processes increased, more efficient methods for monitoring and control were required. This need led to increasing degrees of centralization made possible by developments in measurement, transmission and automatic control. Much less manpower was now required and the operator's role changed to require him to give more attention to supervising the process rather than taking part in the details.

Increasing degrees of centralization introduced more changes. The operator now had more data than he could use effectively. He was also physically separated from the process since most control equipment was placed in a control room. Early control rooms still retained a window to the process unit but even this disappeared with new safety requirements. Consequently, the operator's "touch" with the process was lost. There now developed the man-machine allocation problem. Another important change also occurred. The operator lost his ability to define his own operating procedures. These procedures were now dictated by the control system and display design, and thus by the designer of the system. Tradition in design replaced the operational alternatives previously exercised by the operator.

The traditional operator/process interface today is the panel. It contains the measurement and control information that the operator is assumed to need to carry out his assigned tasks. Because an operator is expected to control a number of plant sections, he must visualize the control panel as a hierarchical display which shows him an overview of the operation as well as allow him to change individual controllers. This hierarchy is shown in Figure 4.

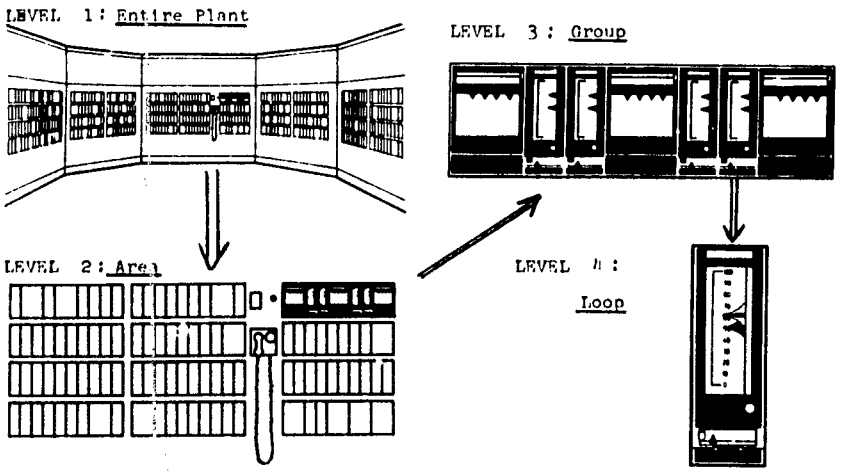


FIGURE 4: CONVENTIONAL PANEL HIERARCHY

Other characteristics of the traditional panel are (Kitchenka, Williamson, 1977):

- Size: Most often the size of a control room is predicated on the size of the panel it must contain. The panel's size is directly proportional to the number of loops it must handle. Considering modern construction standards and expense, many dollars could be saved by reducing the size of this structure.

- Inflexible Layout: Since the panel is usually constructed of steel and each display is wired to suit a particular process control function, changing process control system requirements usually does not result in corresponding changes in display layout.

- Functional Separation: A typical panel is designed with alarm annunciation across the top, analog-type process indication and recording on its vertical midsection, and on-off (motor and electrical) functions on the lower portion. The operator is required to re-associate all of the process-related but physically separated functions necessary to isolate trouble and effectively supervise the process control system.

- Parallel Information Display: The components of a panel present all of the information that they are designed to present all of the time with no potential for selectivity. The operator must, therefore, assimilate this enormous mass of information and attempt to visually and mentally cull out that which is pertinent to any particular situation.

It is important to note that the panel design and display format are usually developed based on the designer's "experience" in understanding the role of the operator and his needs. Infrequently some degree of human factor may be added when the selection of colors or positioning of pushbuttons is involved. In general, the operator display and, at times, the entire control system is almost included as an afterthought to the design of the plant. At times the control system design is even carried out after the process design is fixed (Figure 5) by engineers or contractors who do not have responsibility for the operation of the plant. Consequently, this usually means a study of human operator behavior and a definition of his task was never included in the early planning of the installation. At this late stage, it is economically infeasible to change the process design and, therefore, there is a great tendency to "live with" the control design rather than make changes to the plant design. Consequently, discrepancies between the operational aspects of the control system and the original process objectives can occur. By necessity the changes to the control system design or display format take place after the plant is operating or in the next plant design.

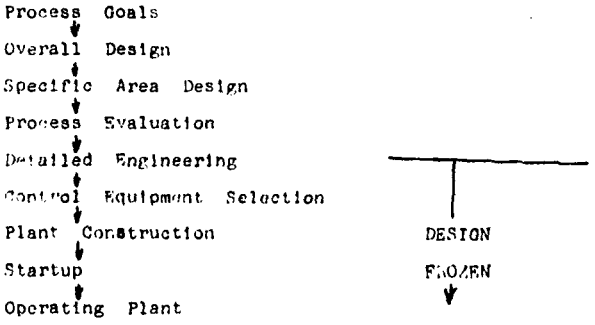


FIGURE 5: THE DESIGN APPROACH FOR PRESENT CONTROL ROOM DISPLAYS

III THE NEXT GENERATION OF MAN-MACHINE INTERFACES DESIGN

The traditional method of operator interface design will have to change in the future to reflect new directions and philosophy in plant operation. Some of the driving forces are:

- Technological changes in electronically driven displays such as CRT's and the advent of microprocessors have increased the pressure to use these devices to solve some of the problems of the panel. Early attempts were not successful, however (Kitchenka, Williamson, 1977).
- Energy and environmental problems are forcing plant managements to require operators to consider other concurrent objectives in addition to maximizing production at specified yield.
- Changes in plant seniority procedures have led to a higher turnover of process operators with resulting change in work load and training levels.
- The introduction of sophisticated processing plants in emerging countries with an unskilled labor force requires changes in the roles the operator and control system play.
- The increased responsibility and disaster potential given operators in terms of the number of control loops and emergency actions are nearing the limits of mental loading (capacity).

In order to successfully meet the challenges presented by these changes, a new methodology for designing operator interfaces will be required in the process industries (Rijnsdorp, Rouse, 1977). This approach will have to consider the functional requirements of the operator first rather than thinking initially in terms of the physical devices (displays, controls, etc.) and letting the operator's procedures be dictated by them. The considerations involve minimizing the human operator's exposure to emergency and stress situations yet giving him a job to perform for which he remains attentive at all times. One may expect that there is an optimum between "too much" and "too little" to do. Additional trade offs between performace and mental load will also have to be made.

To this end, some first actions are already being taken:

- Some companies have included control system and display designs as part of the first phase of process design (Figure 6) which better ensures their harmony with plant operational goals (Beaverstock, 1976).
- The International Purdue Workshop, through its MMI committee, has issued guidelines for man-machine interface design that emphasizes the operator's functional requirements (Purdue University, 1975).
- Some commercial displays have become available that have a functional design as their basis (Kitchenka, Williamson, 1977).
- The problem of acceptance of new technologies and techniques that increase the level of control or management automation are being studied (Beaverstock, Bernard, 1977).

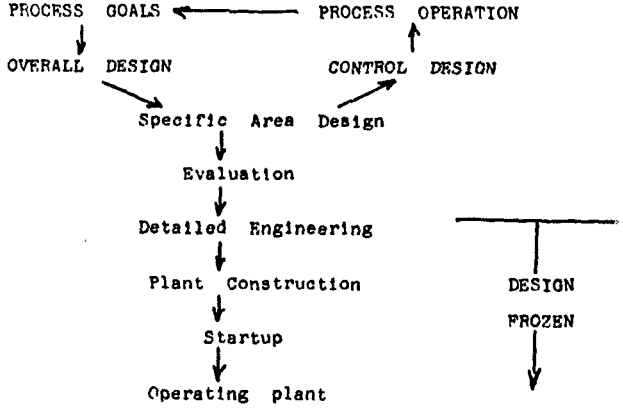


FIGURE 6: THE RECOMMENDED DESIGN APPROACH

These directions are encouraging but still more is required. A much more fundamental approach is necessary. We therefore suggest to model the human operator's behavior in such a way that we can explain and predict to what extent task requirements, system design, disturbances, etc., may affect his performance. This approach follows the many manual control models which have been generated over the past few years for other cases of human supervisory control (Annual Manual Conferences, 1966-77).

In reviewing these we can distinguish two different types of models:

- Descriptive models, i.e., empirically testable models based on some criterion of fit of human input and output data.
- Normative models, i.e., models whose structure is proposed based on a norm or mathematical description and whose parameters are consequently estimated by the human input and output data.

Although most models are of the first type, these models only describe a human performance result - any attempt to explain why a human behaves like the model is not made. An important example of this type is the crossover model based on the describing function method (McRuer, Krendel, 1974). This model however can only be applied to the control of low dimensional systems.

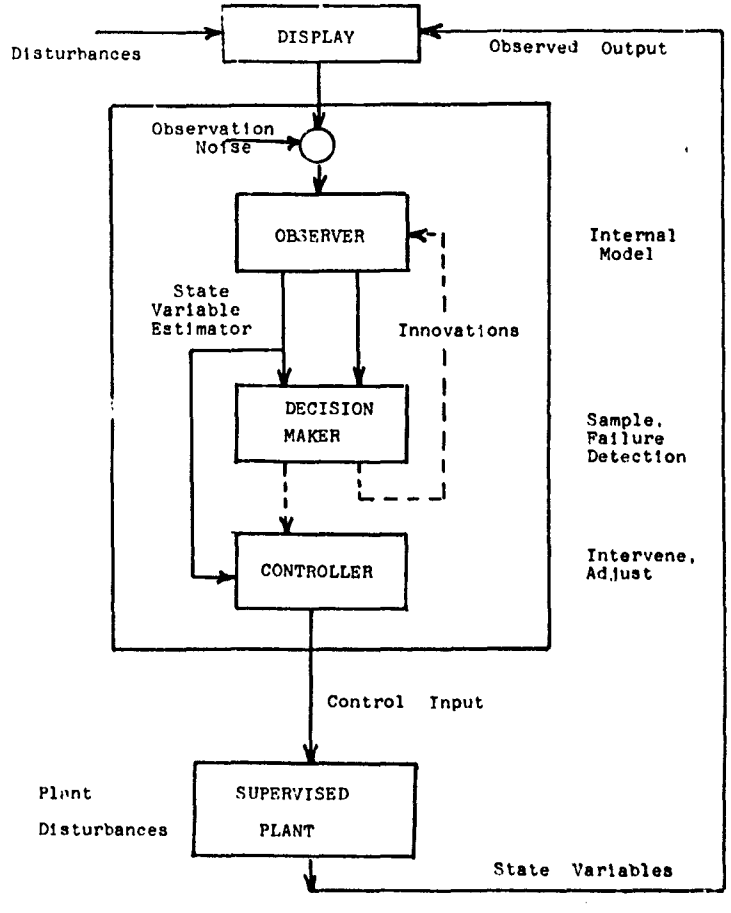
The normative models, however, may be expected to contribute much more to the understanding of the performance of human supervisors since this type of model can be used to verify the action of the human operators. Among the many normative models proposed, such as the decision model (Elkind, 1968) and the model to describe a helmsman's behavior during supertanker control (Veldhuysen, 1976), the most general one is the optimal control model (Baron, Kleinman, 1968). This very general model can also be applied to multivariable, time-varying systems such as found in the process industries.

The model is based on the assumption that a well-trained human controller behaves in an optimal manner taking into account inherent human limitations, constraints, and task requirements. In control engineering terms this means: human beings are optimizing certain criterion. They possess an internal model of the system dynamics, of the system noise and of the task requirement consistent with degree of training or experience they have.

The model mainly consists of two parts (Figure 7): (1) An observer which estimates the minimum variance estimate of the system under control from the displayed information (taking into account observation noise); and (2) A controller and decision-making element that generates setpoints and that decides whether the operator should take over the plant manually due to possible system failures. Hence, the model is capable of describing the following human supervisor functions: monitoring, failure detection, setpoint generators, and manual control or intervening. Through these functions it would be possible to analyze the human operator's ability to control a process plant considering the basic structure of his goals, assigned tasks, knowledge of the process, degree of interaction with the control equipment and the information displayed. Specifically the modeling effort can address: The accuracy of the operator's internal model in making control decisions; the speed at which the operator changes his model under emergency conditions; the relationship between the information displayed and the dynamic reconfigurability of the internal model (including

FIGURE 7

Optimal Control Model of the Human Supervisor



possible saturation effects of too much data); and the relationship between the displays used and the measure of performance. An important feature of the model is the direct relationship between the psychological functions of the human supervisor and the subsections - observer, controller and decision maker - of the model (Kok, vanWijk, 1977).

The task of using the internal model to describe the behavior of a human operator as part of the man-machine system in the process industries is just beginning. The experience in applying the modeling technique is expected to reveal as much about the operator himself as the power of the internal model approach.

IV REFERENCES

1. Annual Conferences on Manual Control Proceedings, Springfield, NTIS.

   2nd NASA - University Conference, Cambridge, Mass. (MIT), 1966, 417 p., NASA SP-126.

   3rd NASA - University Conference, Los Angeles (USC), 1967, 459 p., NASA SP-144.

   4th NASA - University Conference, Ann Arbor (UM), 1968, 594, p., NASA SP-215.

   5th NASA - University Conference, Cambridge, Mass. (MIT), 1969, 713 p., NASA SP-215.

   6th Conference, Dayton (WPAFB), 1970, 296 p.

   7th NASA - University Conference, Los Angeles (USC), 1971, 361 p., NASA SP 281.

   8th Conference, Dayton (WPAFB), 1972, 665 p., AFFDL TR-72-92.

   9th Conference, Cambridge, Mass. (MIT), 1973, 451 p.

   10th Conference, Dayton (WPAFB), 1974, 746 p.

   11th Conference, Moffett Field (ARC), 1975, 711 p., TMX-62, 464.

   12th Conference, Moffett Field (ARC), 1976, 1020 p., TMX-73, 170.

   13th Conference, Cambridge, Mass (MIT), 1977, 400 p.

2. Baron, S., D. L. Kleinman and W. H. Levison, "A Control Theoretic Approach to Manned-Vehicle System Analysis," IEEC Trans. on A.C., Vol. AC-16, No. 6, 1971.

3. Beaverstock, M. C., "The 3 R's of Control System Design," Proceedings of the Third Annual Advanced Control Conference, Purdue University, April 197.

4. Beaverstock, M.C., and J. W. Bernard, "Advanced Control: Ready, Willing, Accepted?"; 5th IFAC/IFID International Conference on Digital Computer Application to Process Control, June 1977.

5. Elkind, J. I. and D. C. Miller, "Adaptive Characteristics of the Human Controller in Time-Varying Systems, " Springfield, NTIS, 1968.

6. Guidelines for the Design of Man-Machine Interfaces for Process Control, MMI Committee of the International Purdue Workshop for Industrial Computer System, Engineering Experiment Station Bulletin, 143 Series, Purdue University, October 1975.

C-1-A

IV REFERENCES (Continued)

7. Kitchenka, F. L. and R. A. Williamson, Jr., "Process Operator Interface Past and Future"; publication pending, Instrumentation Technology.

8. Kok, J.J. and R. A. vanWijk, "A Model of the Human Supervisor, " Annual Conference on Manual Control, MIT, 1977.

9. McRuer, D. J. and E. S. Krendel, "Mathematical Models of Human Pilot Behavior," report; NATO, 1974, AGARD-AG-188.

10. Rijnsdorp, J. E. and W. B. Rouse, "Design of Man-Machine Interfaces in Process Control," 5th IFAC/IFIP Conference, The Hague - The Netherlands, June, 1977.

11. Veldhuyzen, W. and H. G. Stassen, "The Internal Model: A Useful Concept in the Analysis of Man-Machine Systems."